

# Degrees of Freedom of the Bursty MIMO X Channel without Feedback

Shih-Yi Yeh and I-Hsiang Wang

## Abstract

We study the sum degrees of freedom (DoF) of the bursty MIMO X channel without feedback, where the four transmitter-receiver links are intermittently on-and-off, controlled by four Bernoulli random sequences which may be arbitrarily correlated, subject to a symmetry assumption: The two direct-links have the same level of burstiness, modeled by  $\text{Ber}(p_d)$ , and so do the cross-links, modeled by  $\text{Ber}(p_c)$ . The sum DoF is fully characterized in the regime where  $\frac{p_c}{p_d}$  is small, i.e. below a certain threshold, and is partially characterized in the other regime where  $\frac{p_c}{p_d}$  is above the threshold. The achievability is proved with a combination of Han-Kobayashi strategy and interference alignment, which can achieve strictly higher DoF than interference alignment alone. The converse proof employs a channel-state-sequence pairing technique. Striking differences are also found between the bursty and the non-bursty X channel. In particular, various interference alignment schemes that achieve the DoF of non-bursty X channels are found to be suboptimal when the channels become bursty. The reciprocity between the forward and the reverse links is lost, and the sum DoF does not saturate when the ratio of the transmitter and the receiver antennas exceeds  $\frac{2}{3}$ .

## I. INTRODUCTION

With the rapid deployment of wireless networks and the ever-increasing number of devices connected to them, interference becomes more and more prevalent and is often one of the key factors limiting the capacity of a network. For example, the sum degrees of freedom (DoF) of a two-user-pair MIMO interference channel (IC) with  $M$  antennas at all nodes is only  $M$  [1], instead of  $2M$  as it would be without interference. Allowing cross-link messaging (i.e. Tx1 to Rx2 and Tx2 to Rx1 as in Fig. 1) in such an interference network as proposed in [2] may serve as a simple remedy. It boosts the sum DoF from  $M$  to  $4M/3$ , and quite remarkably this gain can be realized with simple interference alignment (IA) schemes. The two-user-pair IC with cross-link messaging allowed is coined X channel (XC) in the literature, and its DoF is characterized in [3] and [4].

However, the DoF characterization of the MIMO X channel above assumes that the four transmitter-receiver (Tx-Rx) links are always present throughout the transmission, which neglects some phenomenons that may be important in real scenarios. For example, fading is a critical issue for wireless networks and will result in time-varying channel

The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Barcelona, Spain, July 2016.

S.-Y. Yeh is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan (email: steven0416@gmail.com).

I.-H. Wang is with the Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (email: ihwang@ntu.edu.tw).

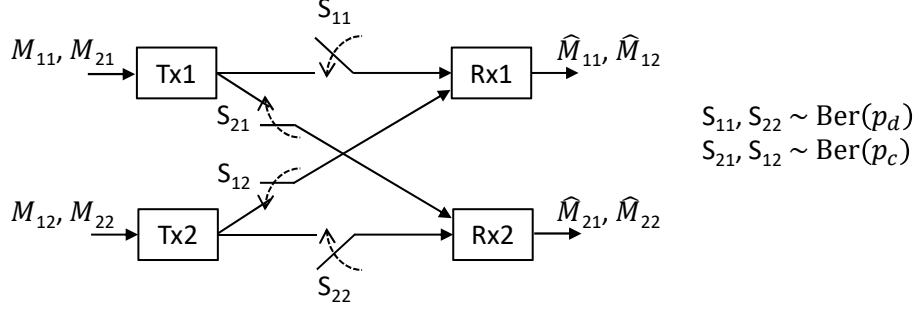


Fig. 1. Bursty MIMO X Channel ( $M_{ji}$ : message transmitted from  $Tx_i$  to  $Rx_j$ ,  $\hat{M}_{ji}$ : decoded message)

matrices for the four Tx-Rx links. As a first step toward understanding the fading effects on the channel capacity, we may consider the shadowing phenomenon and model it with four randomly on-and-off Tx-Rx links. This may also serve as a simple abstraction of the co-channel interference effects when there are other wireless networks operated in the vicinity ([5], [6]), which can be more commonplace in the era of internet of things (IoT). Machine-type communications envisioned for the IoT era can also lead to short bursts of interruption to the regular MIMO X channel traffic, and their effects can be captured with four intermittent Tx-Rx links too. Lastly, if frequency hopping is allowed, the two pairs of users may hop to two different frequency bands and enjoy higher sum rate, even without multi-carrier capability in the physical layer. In this case, the cross-links may be intermittently off.

This motivates us to investigate the fundamental limit of the *bursty* MIMO X channel, where each of the four Tx-Rx links can be intermittently on-and-off as illustrated in Fig. 1 ( $S_{ij} = 1$  : the link is on,  $i, j \in \{1, 2\}$ .) The burstiness of the four Tx-Rx links is governed by four Bernoulli random sequences which can be arbitrarily correlated at each time instant, subject to the following channel symmetry assumption: The direct-links (i.e. Tx1 to Rx1 and Tx2 to Rx2) have the same level of burstiness, modeled by  $\text{Ber}(p_d)$ , and so do the cross-links, modeled by  $\text{Ber}(p_c)$ , and the direct-links also have the same level of *conditional burstiness*, i.e.  $\mathcal{P}(S_{11} = 1 | S_{12} = 1) = \mathcal{P}(S_{22} = 1 | S_{21} = 1) = p_{d|c}$ . As we will see in later sections, this symmetry simplifies the analysis of the DoF. Note that this model is very similar to the one in [6], except that correlation among the four Bernoulli random sequences are allowed and it is applied to the MIMO X channel where each transmitter is equipped with  $M$  antennas and each receiver with  $N$  antennas. We study the sum DoF of this bursty MIMO X channel without feedback. (For clarity, we shall refer to the usual MIMO X channel whose links are always present as the *non-bursty* channel.)

Our main contribution is the characterization of the sum DoF of the bursty MIMO X channel when its  $M, N, p_c, p_d$  and  $p_{d|c}$  parameters fall in the solid regions of Fig. 2. The DoF in the dotted regions remains unknown. The channels are categorized into two types according to Tx and Rx antenna numbers - *Type I* has  $M \leq N$  and *Type II* has  $N \leq M$ . For each channel type, we distinguish two operating regimes, similar to [6]. The channel is said to be in *Regime 1* when  $p_c/p_d \leq 1/(1 + p_{d|c})$ , and its sum DoF is fully characterized in this regime. On the other hand, if  $1/(1 + p_{d|c}) < p_c/p_d \leq 1$ , then the channel is operated in *Regime 2*, where the sum DoF is only partially known, and as can be seen in Fig. 2, we know a little more about the DoF of Type II than Type I. (It will become clear

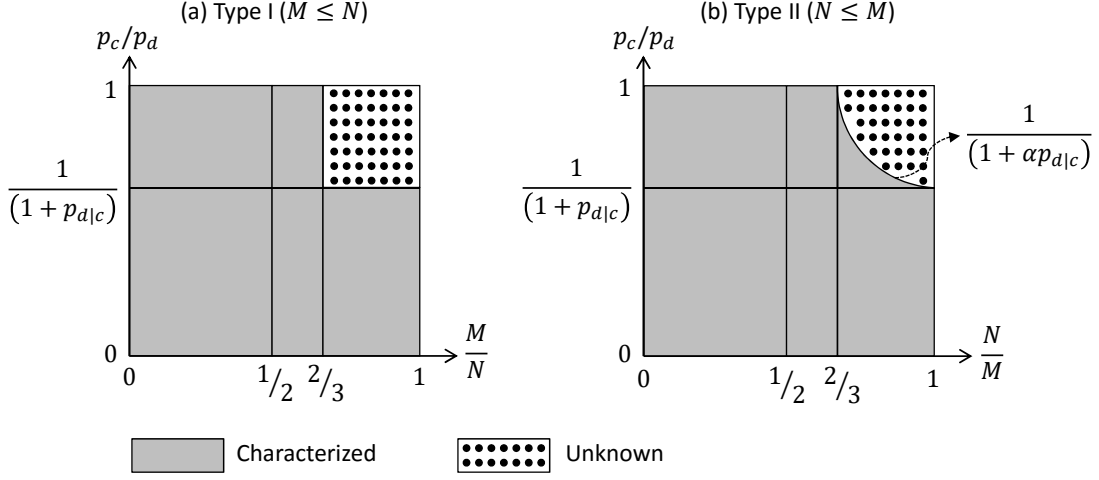


Fig. 2. Status of sum DoF characterization ( $\alpha = \frac{3N-2M}{2N-M}$ )

shortly that it suffices to characterize the channel for  $p_c \leq p_d$  due to channel symmetry.) The sum DoF ( $\eta$ ) of the channel in Regime 1 and 2 is illustrated in Fig. 3(a) and 3(b), respectively.

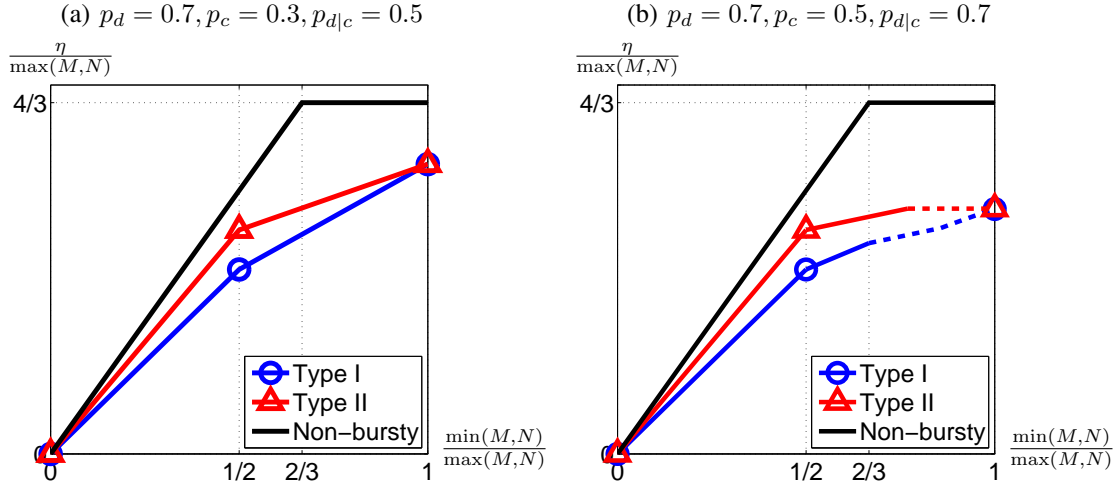


Fig. 3. Example DoF plots (Dashed lines: lower bounds.)

Through the study of the DoF of the bursty X channel, we also found striking differences compared to the non-bursty X channel. For example, various IA schemes that achieve the DoF of non-bursty MIMO X channels are found to be sub-optimal when the channels become bursty. New coding schemes are needed for these channels even when the burstiness of the channel (CSI) is not known to the transmitters. In particular, we discovered a combination of the Han-Kobayashi (HK) scheme and IA scheme that can outperform IA for the bursty MIMO X channels. In addition, the non-bursty MIMO X channels are known to have the following properties:

- (i) The Type I and the Type II channel are reciprocal, i.e. if we swap  $M$  and  $N$ , the sum DoF is unchanged.

- (ii) For the Type I channel ( $M \leq N$ ), the sum DoF saturates at  $4N/3$  when  $M$  grows beyond  $2N/3$ . Similarly the DoF of the Type II channel saturates at  $4M/3$  when  $N$  reaches  $2M/3$ .
- (iii) The presence of cross-links does not increase the sum DoF of the channel.

However, when the channels become bursty, *none* of these are true in general. The contrasts to (i) and (ii) are evident in Fig. 3. Other contrasts will be shown in later paragraphs. These sharp contrasts stem from the fact that burstiness creates new channel states, which are under-utilized by the DoF-achieving IA schemes of the non-bursty X channel.

For the converse proof of the DoF, we extend the state-sequence pairing technique developed in [7] and derive a DoF upper bound. However, unlike [7], where the channel burstiness is controlled by a single Bernoulli random sequence, our channel now has four Bernoulli random sequences. So the converse proof becomes much more complicated, as we have 16 channel states at each time instant, as opposed to only two previously. This difficulty, however, can be overcome with the so-called “contracted channel” state reduction technique in [6], which reduces the channel states to only five. We then use Gaussian distributions to bound a number of differential entropies, and arrive at a DoF upper bound that is tight everywhere in the solid regions of Fig. 2.

The DoF achievability is proved with schemes that are motivated from the linear IA schemes ([4], [3]) and the HK scheme ([6], [9], [8]). The linear IA schemes in [4] can be called *interference-nulling beamforming (INBF)* schemes, as they use beamforming to null out the interference, thereby orthogonalizing the channel. INBF schemes suit our needs particularly well. However, due to the new channel states induced by the burstiness of the channel, more sophisticated INBF schemes are needed to achieve the DoF in many cases. Furthermore, for the solid region of Fig. 2(b) where  $1/2 < N/M \leq 1$  and  $p_c/p_d > 1/(1 + p_{d|c})$ , the INBF schemes no longer achieve the DoF of the channel. In this case, superimposing public messages on top of the INBF private messages and decoding them successively similar to the celebrated HK scheme in [9] can yield strictly higher DoF than using the INBF alone, and in fact achieve the DoF of the channel. For brevity, we name this scheme *HKIA*, since it combines the HK strategy and the IA scheme. As for the dotted regions of the Fig. 2, the achievable DoF of our HKIA scheme does not coincide with the upper bound, so the DoF remains unknown there. However, the HKIA can again achieve higher DoF than the INBF, so the INBF schemes are not DoF-optimal for these channel configurations either.

### Related Works

This paper is concerned with the DoF of the bursty MIMO X channel. To put this work into better perspective, it helps to have a brief review of the works on the bursty interference channel, in addition to the works on the MIMO X channel [2]–[4], which we have referred to at the beginning of this section. First of all, note that although the interference channel (IC) has been studied since the 1970s [10], the bursty IC received little attention until recently. [11] and [12] studied the two-user-pair IC where interference (cross-links) may be absent with a degraded message set formulation. [7] and [13] considered the capacity of the bursty IC when delayed feedback is available. A similar delayed-feedback problem was treated in [14], where both direct-links and cross-links can be intermittently on-and-off in a binary fading IC. [6] then investigated the symmetric binary fading IC with no feedback in the

system. Lastly, a simplified version of our work was presented in [15], which is the first to consider the DoF of the bursty MIMO X channel without feedback. It assumes a simple channel model where only the cross-links can be intermittent, controlled by a single Bernoulli random sequence.

The rest of this paper is structured as follows. The formal problem formulation is given in Section II. The main results are summarized in Section III. The fundamental differences of the bursty MIMO X channel, compared to the non-bursty channel, are presented in Section IV. We then derive the sum DoF upper bound in Section V, and show its achievability in Section VI. Section VII provides a lower bound of the sum DoF where the DoF remains unknown. Section VIII addresses some open and subtle issues. Finally, we conclude this paper with the remarks in Section IX.

## II. PROBLEM FORMULATION

The detailed system model of the bursty MIMO X channel considered in this work is depicted in Figure 4. There are two transmitters and two receivers in the system, denoted by  $\text{Tx}i$  and  $\text{Rx}j$ , respectively, for  $i, j \in \{1, 2\}$ . Each transmitter is equipped with  $M$  antennas, while each receiver has  $N$  antennas.  $M_{ji} \sim \text{Unif}\{1, 2, \dots, 2^{nR_{ji}}\}$  denotes the message from  $\text{Tx}i$  to  $\text{Rx}j$ , with rate  $R_{ji}$ , and  $\hat{M}_{ji}$  is the decoded message at  $\text{Rx}j$ .  $X_i$  represents the signal transmitted by  $\text{Tx}i$  and  $Y_j$  is the received signal at  $\text{Rx}j$ . Each transmitter has an average transmit power constraint  $P$ , i.e.  $\frac{1}{n} \sum_{k=1}^n \|X_i[k]\|^2 \leq P$ ,  $i \in \{1, 2\}$ , where  $X_i[k]$  denotes the  $k$ -th transmitted symbol of  $\text{Tx}i$ .  $H_{ji}$  models the  $N \times M$  channel matrix from  $\text{Tx}i$  to  $\text{Rx}j$ . The channel matrices are drawn randomly from a continuous distribution with i.i.d. elements, but are fixed during the transmission. Each transmitter or receiver is assumed to have perfect knowledge of all channel matrices.  $Z_j$  is the additive Gaussian noise at  $\text{Rx}j$  with zero mean and unit variance, i.i.d. in time.

The four Tx-Rx links can be intermittently on and off, controlled by four Bernoulli random sequences,  $S_{11}[k]$ ,  $S_{12}[k]$ ,  $S_{21}[k]$ , and  $S_{22}[k]$ , which can be arbitrarily correlated at each time instant and are i.i.d. in time. The link from  $\text{Tx}i$  to  $\text{Rx}j$  is on at the  $k$ -th time instant when  $S_{ji}[k] = 1$ , and it is off when  $S_{ji}[k] = 0$ . For brevity of notation, we may drop the dummy time index ( $k$ ) herein and abbreviate  $S_{ji}[k]$  as  $S_{ji}$  when there is no confusion. As mentioned

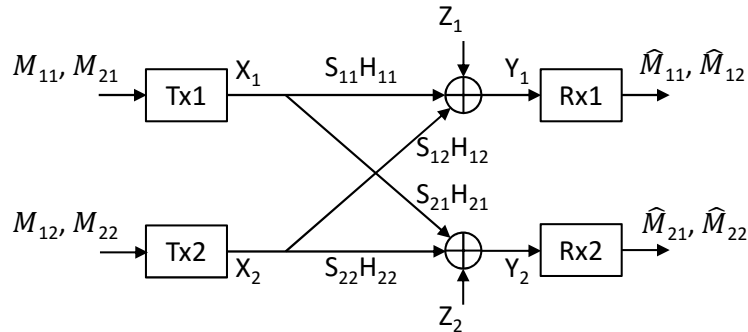


Fig. 4. Model of the bursty MIMO X channel

earlier, in this paper we consider the symmetric case where  $(S_{11}, S_{12})$  has the same distribution as  $(S_{21}, S_{22})$ , with  $S_{11} \sim \text{Ber}(p_d)$ ,  $S_{12} \sim \text{Ber}(p_c)$ ,  $\mathcal{P}(S_{11} = 1 | S_{12} = 1) = p_{d|c}$ , and  $\mathcal{P}(S_{11} = 1, S_{12} = 1) = p_{cd} = p_c p_{d|c}$ . Each receiver has perfect knowledge of the current burstiness of the two incoming links, e.g. Rx1 knows  $S_{11}$  and  $S_{12}$ . There is no feedback in the system and transmitters do not know any  $S_{ij}$  variables.

Let  $R_{ji}$  be the code rate from Tx $i$  to Rx $j$ ,  $i, j \in \{1, 2\}$ . A rate tuple  $(R_{11}, R_{12}, R_{21}, R_{22})$  is said to be achievable on the bursty MIMO X channel if there exists a sequence of codes such that  $\mathcal{P}\{\hat{M}_{ji} \neq M_{ji}, \text{ for some } i, j \in \{1, 2\}\}$  converges to zero as the block length of the codes tends to infinity. The capacity region of the channel is the set of all achievable rate tuples  $(R_{11}, R_{12}, R_{21}, R_{22})$ . The sum DoF,  $\eta$ , of the channel follows conventional definition, i.e.

$$\eta \triangleq \lim_{P \rightarrow \infty} \frac{C_{\text{sum}}}{(\frac{1}{2}) \log(P)}, \quad (1)$$

where  $C_{\text{sum}}$  is the sum capacity of the channel. Note that in this paper we evaluate the sum DoF of the channel in the almost surely (a.s.) sense, since the channel matrices are drawn from a continuous probability distribution as in [3].

The following notations are used herein to simplify the DoF derivations and discussions.

- (i)  $\langle \eta_{cd}, \eta_{\bar{c}d}, \eta_{c\bar{d}} \rangle$ : For the DoF analysis in this paper, it is often natural or convenient to express the DoF as a linear combination of  $p_{cd}$ ,  $p_d - p_{cd}$ , and  $p_c - p_{cd}$ , which are the probabilities of  $(S_{ij}, S_{ii})$  being  $(1, 1)$ ,  $(0, 1)$ , and  $(1, 0)$ , respectively,  $i \neq j$ . So we define  $p_{\bar{c}d} \triangleq p_d - p_{cd}$ , and  $p_{c\bar{d}} \triangleq p_c - p_{cd}$ , and introduce  $\langle \eta_{cd}, \eta_{\bar{c}d}, \eta_{c\bar{d}} \rangle$ , or the  $\eta$ -triple, to denote  $\eta_{cd}p_{cd} + \eta_{\bar{c}d}p_{\bar{c}d} + \eta_{c\bar{d}}p_{c\bar{d}}$ . For example  $\langle 2, 1, 1 \rangle = 2p_{cd} + 1p_{\bar{c}d} + 1p_{c\bar{d}}$ . Note that the  $\eta$ -triple is a *scalar* value, not a vector. When it helps, we may also add subscript  $\mathcal{P}_S$  to the  $\eta$ -triple (i.e.  $\langle \eta_{cd}, \eta_{\bar{c}d}, \eta_{c\bar{d}} \rangle_{\mathcal{P}_S}$ ) to emphasize its dependence on the distribution of the state variables,  $\{S_{ij} \mid i, j \in \{1, 2\}\}$ .
- (ii)  $\mathcal{D}\{x\}$ : Quite often the DoF computation involves entropies and mutual informations, so we define  $\mathcal{D}\{x\} \triangleq \lim_{P \rightarrow \infty} \frac{x}{(1/2) \log(P)}$ , and use it in places like  $\mathcal{D}\{h(Y_1 \mid S_{11}, S_{12})\}$ ,  $\mathcal{D}\{I(X_1; Y_1 \mid S_{11}, S_{12})\}$ , et al.
- (iii)  $X^n$ : We adopt the convention of using  $X^n$  to represent a sequence (vector) of random variables,  $X[1], X[2], \dots, X[n]$ , where the number inside  $[\ ]$  is the time index. For instance,  $Y_1^n$  denotes the random vector consisting of  $n$  received symbols at Rx1, i.e.  $Y_1[1], Y_1[2], \dots, Y_1[n]$ . Similarly,  $(S_{11}, S_{12})^n$  indicates the following set of derived random variables  $\{(S_{11}[k], S_{12}[k]) \mid k = 1, 2, \dots, n\}$ , and  $(S_{12}H_{12}X_2 + Z_1)^n = \{S_{12}[k]H_{12}X_2[k] + Z_1[k] \mid k = 1, 2, \dots, n\}$ . (Recall that the channel matrices are fixed throughout the transmission.)
- (iv)  $M \times N$  channel: We refer to a MIMO channel with  $M$  antennas at the transmitter and  $N$  antennas at the receiver as an  $M \times N$  channel, which can be bursty or non-bursty.

### III. MAIN RESULTS

We summarize our main results in this section. First of all, as mentioned in Section I, the sum DoF characterization status of the bursty MIMO X channel is shown in Fig. 2, where the DoF is characterized in the solid regions and remains unknown in the dotted regions. The sum DoF in the solid regions is established by Theorem 1 and 2.

*Theorem 1:* The sum DoF of the bursty MIMO X channel is upper bounded by  $\eta_{\text{ub}}$  (a.s.), with

$$\eta_{\text{ub}} \triangleq \begin{cases} 2[p_{cd} \min(2M, N) + p_c \min(M, 2N) + (p_d - p_c - 2p_{cd}) \min(M, N)], & \frac{p_c}{p_d} \leq \frac{1}{1 + p_{d|c}} \\ 2[(p_d - p_c) \min(2M, N) + (p_d - p_{cd}) \min(M, 2N) + (p_{cd} - 2p_d + 2p_c) \min(M, N)], & \frac{p_c}{p_d} > \frac{1}{1 + p_{d|c}}. \end{cases}$$

*Theorem 2:* The upper bound,  $\eta_{\text{ub}}$ , in Theorem 1 is tight except when

- (i)  $2/3 < M/N \leq 1$  and  $\frac{1}{1+p_{d|c}} < \frac{p_c}{p_d} \leq 1$ , or
- (ii)  $2/3 < N/M \leq 1$  and  $\frac{1}{1+\alpha p_{d|c}} < \frac{p_c}{p_d} \leq 1$ ,

where  $\alpha \triangleq \frac{3N-2M}{2N-M}$ .

Note that  $0 < \alpha \leq 1$ , when  $2/3 < N/M \leq 1$ , so the region where the DoF is unknown is smaller for Type II than for Type I. When  $\eta_{\text{ub}}$  is not tight, the following theorem provides a lower bound of the sum DoF, which can be strictly higher than the DoF achieved by the INBF scheme. This lower bound is more conveniently expressed in the  $\eta$ -triple notation.

*Theorem 3:* If  $2/3 < M/N \leq 1$  and  $1/(1 + p_{d|c}) < p_c/p_d \leq 1$ , then, with  $N = 3k + q$  where  $k$  is an integer and  $q \in \{0, 1, 2\}$ , the sum DoF is lower bounded by  $\eta_{\text{lb},1}$ , defined as

$$\eta_{\text{lb},1} \triangleq \begin{cases} \langle 4k + q, 2k + q, 2M - 2k - q \rangle, & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_1 p_{d|c}} \\ \langle N, M, M \rangle + \left\langle k, \left\lfloor \frac{q}{2} \right\rfloor, -\left\lfloor \frac{q}{2} \right\rfloor \right\rangle \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - \lceil \frac{q}{2} \rceil, 2k + \lceil \frac{q}{2} \rceil \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_1 p_{d|c}}, \end{cases}$$

where  $\beta_1 \triangleq \frac{2M-4k-q}{2k+q}$ . On the other hand, if  $2/3 < N/M \leq 1$  and  $1/(1 + \alpha p_{d|c}) < p_c/p_d \leq 1$ , then, with  $M = 3k + q$  where  $k$  is an integer and  $q \in \{0, 1, 2\}$ , the sum DoF is lower bounded by  $\eta_{\text{lb},2}$ , defined as

$$\eta_{\text{lb},2} \triangleq \begin{cases} M \left( \langle 1, 1, 1 \rangle + \langle 1, 0, 0 \rangle \frac{\langle 1, -1, 1 \rangle}{\langle 3, -2, 2 \rangle} \right), & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_2 p_{d|c}} \\ M \langle 1, 1, 1 \rangle + k \langle 1, 0, 0 \rangle, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_2 p_{d|c}}, \end{cases}$$

where  $\beta_2 \triangleq \frac{q}{k+q}$ .

Sharp contrasts exist between the DoF of the bursty and non-bursty MIMO X channels, including those listed in Table I. (For C1,  $\max(M, N)$  is assumed to be a multiple of 3 when  $\frac{\min(M, N)}{\max(M, N)} > \frac{2}{3}$  and  $M \neq N$ .)

Table I  
CONTRASTS BETWEEN BURSTY AND NON-BURSTY MIMO X CHANNELS

| Contrast | Description  | Non-bursty | Bursty |
|----------|--|------------|--------|
| C1       | Does IA always achieve the optimal DoF?                          | Yes        | No     |
| C2       | Are Type I and Type II reciprocal?                               | Yes        | No     |
| C3       | Does the DoF saturate at $\frac{\min(M, N)}{\max(M, N)} = 2/3$ ? | Yes        | No     |
| C4       | Can crosslinks increase the DoF?                                 | No         | Yes    |

#### IV. FUNDAMENTAL DIFFERENCES

In this section we look into the fundamental differences of the bursty MIMO X channel, compared to the non-bursty channel, to bring out the intuition of the contrasts between them. We do so with the help of a few concrete and simple channels.

##### A. IA Does Not Always Achieve the Optimal DoF

The first and foremost difference between the bursty and non-bursty MIMO X channels is that burstiness of the channel creates new channel states, which may be under-utilized by the IA schemes that achieve the DoF of the non-bursty channel. To see this, let us consider the simple  $3 \times 2$  MIMO X channel ( $M = 3, N = 2$ ). When this channel is non-bursty, it is well-known that the sum DoF is 4 (a.s.), and can be achieved by the INBF scheme shown in Fig. 5, which decomposes the X channel into two orthogonal multiple-access channels (MACs) with the four  $3 \times 1$  beamforming unit vectors satisfying  $H_{21}\psi_1 = H_{11}\phi_1 = H_{22}\phi_2 = H_{12}\psi_2 = 0$ . When the channel becomes bursty, it is easily verified that the DoF achieved by this scheme is  $2(p_c + p_d)$ , or  $2\langle 2, 1, 1 \rangle$  (a.s.). The sub-optimality of this scheme is evident upon an examination of the received signal at Rx1. (The situation at Rx2 is the same.) Specifically, consider the channel states for Rx1 illustrated in Fig. 6, where a solid line indicates that the link is on, a dashed line indicates an off link and each state is named by its probability,  $p_{cd}, p_{\bar{c}d}$ , or  $p_{c\bar{d}}$ . The dimensions of the received signal at Rx1 in these three channel states are 2, 1, and 1 (a.s.), respectively. However, we have two antennas at Rx1 and three antennas at each transmitter, so it is clear that we under-utilize the available dimensions of the  $p_{\bar{c}d}$  and  $p_{c\bar{d}}$  channel states. (Note that we omit the fourth channel state where both links to Rx1 are off and no data is communicated.)

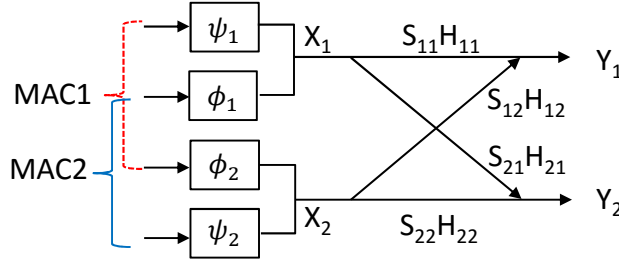


Fig. 5. The INBF scheme for the  $3 \times 2$  channel

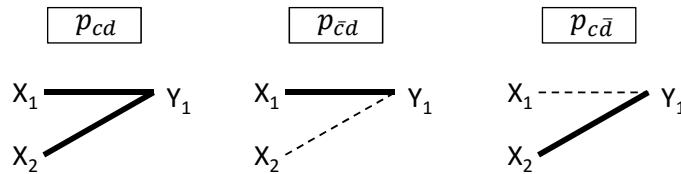


Fig. 6. The channel states for Rx1 (The states for Rx2 are defined symmetrically.)



This suggests that by taking advantage of the unused dimensions in the  $p_{\bar{c}d}$  and  $p_{c\bar{d}}$  channel state, we may achieve higher DoF. One possible way is to introduce public messages. Let  $M_{01}$  and  $M_{02}$  be the public messages sent from Tx1 and Tx2, respectively, to be decoded at both receivers, and consider the following random coding scheme:

$$X_1^n(M_{01}, M_{11}, M_{21}) = U_1^n(M_{01}) + [\psi_1 \ \phi_1] \begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix},$$

where  $U_1 \sim \mathcal{N}(0, \frac{P}{6} I_3)$ ,  $C_1 \& D_1 \sim \mathcal{N}(0, \frac{P}{4})$ , and  $[\psi_1 \ \phi_1] \begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}$  is a shorthand notation for multiplying

each of the  $n$  components of  $\begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}$  by  $[\psi_1 \ \phi_1]$ .  $X_2^n(M_{02}, M_{12}, M_{22})$  is encoded symmetrically. At each receiver, we decode the public messages first, remove them, and then decode the private messages, similar to the HK scheme in [9]. With this decoding order, clearly the sum DoF achieved by the private messages remains unchanged, i.e.  $2\langle 2, 1, 1 \rangle$ . So the key question is: What DoF is achievable on the public messages? Since the public messages are decoded at two MACs, it follows that  $\eta_{01}$  and  $\eta_{02}$  (DoF) satisfying the following inequalities are achievable on  $M_{01}$  and  $M_{02}$ , respectively:

$$\eta_{01} \leq \mathcal{D}\{\min[I(U_1; Y_1 | S, U_2), I(U_1; Y_2 | S, U_2)]\}$$

$$\eta_{02} \leq \mathcal{D}\{\min[I(U_2; Y_1 | S, U_1), I(U_2; Y_2 | S, U_1)]\}$$

$$\eta_{01} + \eta_{02} \leq \mathcal{D}\{\min[I(U_1, U_2; Y_1 | S), I(U_1, U_2; Y_2 | S)]\}.$$

Following the steps in Section VI, it can be shown that  $\eta_{01} = \eta_{02} = \langle 0, 0, 1 \rangle$ , or  $p_c - p_{cd}$ , is achievable, which is *non-zero* whenever  $p_{cd} < p_c$ ! The plausibility of this result can also be seen intuitively as follows. Recall that the dimensions of the private messages at each receiver are 2, 1, and 1 in the three channel states (Fig. 6), so  $\eta_{01} + \eta_{02}$  is upper bounded by  $\langle 0, 1, 1 \rangle$ , since there is one free dimension only in the 2nd and 3rd states. However, each public message also has to be decoded at both receivers. For  $M_{01}$  to be decoded at Rx1, the  $\eta_{01}$  cannot be greater than  $\langle 0, 1, 0 \rangle$ , because the link from Tx1 to Rx1 exists only in the first two states. Similarly, for  $M_{01}$  to be decoded at Rx2, we need  $\eta_{01} \leq \langle 0, 0, 1 \rangle$ . Since  $p_c \leq p_d$ , it follows that  $\langle 0, 0, 1 \rangle \leq \langle 0, 1, 0 \rangle$ , and hence  $\eta_{01} \leq \langle 0, 0, 1 \rangle$ . Finally, noting that the DoF of  $M_{02}$  is symmetric to that of  $M_{01}$ , we conclude that  $\eta_{01} = \eta_{02} = \langle 0, 0, 1 \rangle$  is achievable, and the total DoF achievable is  $2\langle 2, 1, 2 \rangle$ . And this is in fact the sum DoF of this channel! This coding scheme is named HKIA, as it incorporates the ideas of the HK strategy and the IA scheme. The DoF achieved by the INBF (IA) and the HKIA schemes on this channel ( $N/M = 2/3$ ) is illustrated in Fig. 7.

So, we have seen that for the  $3 \times 2$  bursty MIMO X channel, the earlier INBF scheme no longer achieves the DoF of the channel, because the  $p_{\bar{c}d}$  and  $p_{c\bar{d}}$  states are under-utilized. The HKIA scheme takes advantage of them by superimposing public messages on the private messages and achieves the DoF of the channel.

### B. Type I and II Are Not Reciprocal

The loss of reciprocity between Type I and Type II bursty channels has been seen in Fig. 3, and *some* intuition can be gained by considering two simple channels, the  $1 \times 3$  and  $3 \times 1$  bursty X channels. The DoF of the bursty

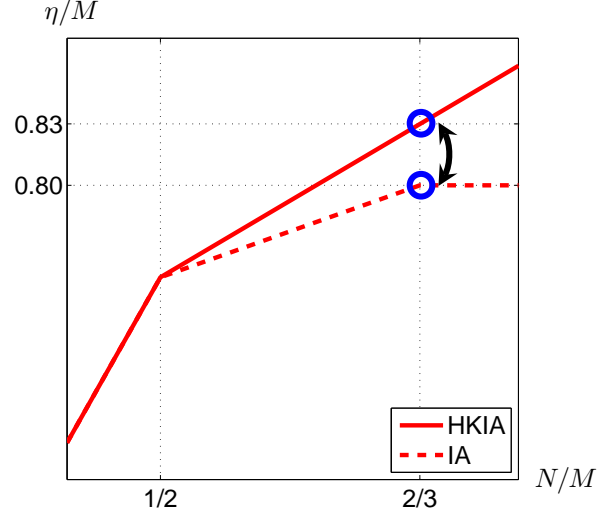


Fig. 7. DoF achieved by IA and HKIA on the  $3 \times 2$  bursty channel with  $p_d = 0.7$ ,  $p_c = 0.5$ , and  $p_{d|c} = 0.9$

$3 \times 1$  channel can be shown to be  $2(p_d + p_c - p_{cd})$ , or  $2\langle 1, 1, 1 \rangle$ , and is achievable by the exactly same INBF scheme as shown in Fig. 5, since each effective MAC clearly achieves a DoF of  $p_d + p_c - p_{cd}$ .

However, if we consider the reciprocal INBF scheme for the  $1 \times 3$  bursty X channel as depicted in Fig. 8, does it achieve the same DoF? The answer is *no*, and can be seen as follows. This INBF scheme decomposes the  $1 \times 3$  bursty X channel into two orthogonal broadcast channels (BCs) with the four  $3 \times 1$  beamforming vectors satisfying  $\psi_1^* H_{12} = \phi_1^* H_{11} = \phi_2^* H_{22} = \psi_2^* H_{21} = 0$ . So each effective broadcast channel is a  $1 \times 1$  broadcast channel, and its DoF is the same as that of a *degraded*  $1 \times 1$  broadcast channel, i.e.  $p_d$ , following similar lines of arguments in [1]. Consequently this INBF scheme achieves only  $2p_d$  DoF, which is in fact the DoF of the bursty  $1 \times 3$  channel and is strictly less than  $2(p_d + p_c - p_{cd})$  whenever  $p_c > p_{cd}$ .

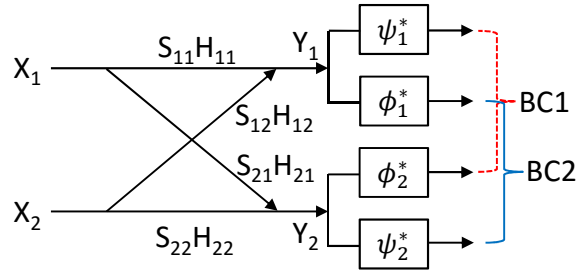


Fig. 8. The INBF scheme for the  $1 \times 3$  channel

Therefore in light of the fundamental difference between the bursty BC and MAC, it is understandable that bursty Type I and Type II channels are not reciprocal in general, and a bursty Type II channel can have higher DoF than the corresponding Type I channel.

### C. Cross-links Can Increase the DoF

The bursty  $3 \times 1$  channel considered in the previous subsection also sheds light on why the presence of the cross-links can increase the DoF of the MIMO X channel when the channel becomes bursty. Recall that with the INBF scheme in Fig. 5, the DoF of each MAC is  $p_d + p_c - p_{cd}$ , or  $\langle 1, 1, 1 \rangle$ . On the other hand, in the absence of the cross-links (i.e.  $p_c = 0$ ), the channel reduces to two orthogonal  $3 \times 1$  point-to-point bursty MISO channels, each of which simply has a DoF of  $p_d$ , or  $\langle 1, 1, 0 \rangle$ . Therefore we see that the cross-links increases the DoF of the bursty  $3 \times 1$  X channel because of the presence of the  $p_{c\bar{d}}$  state at each MAC. Since  $p_{c\bar{d}} = p_c - p_{cd}$  and  $p_{cd} = p_c$  when  $p_d = 1$ , this state is suppressed for the non-bursty MIMO X channel, so the presence of the cross-links does not help the DoF. Only when the channel becomes bursty, can the  $p_{c\bar{d}}$  state emerge and present an opportunity for data transmission, leading to higher DoF. This explains why cross-links can increase the DoF for the bursty MIMO X channels, but not for the non-bursty channels.

### D. DoF Does Not Saturate at $\frac{\min(M,N)}{\max(M,N)} = \frac{2}{3}$

An easy way to see why the sum DoF does not saturate when  $\frac{\min(M,N)}{\max(M,N)} = \frac{2}{3}$  for the bursty MIMO X channel in general (c.f. Fig. 3) is to consider a Type I channel ( $M \leq N$ ) operated in Regime 1 with small  $p_c$  and large  $p_d$ . (The case of Type II is similar.) Examining the channel states for Rx1 and Rx2 (c.f. Fig. 6), we conclude that the probability of the  $p_{c\bar{d}}$  state will be large in this scenario, while all other states will have small probabilities. This implies that we essentially have two orthogonal point-to-point  $M \times N$  channels for a large fraction of time. So intuitively the DoF should keep growing as  $M$  increases, since  $M \leq N$ . Two bursty Type I DoF curves are illustrated in Fig. 9.

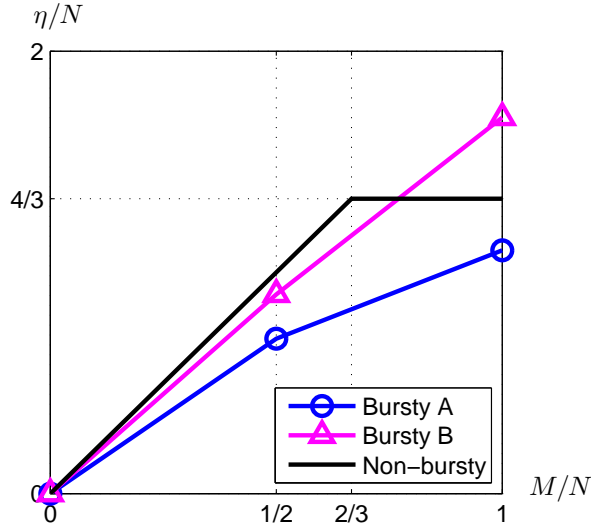


Fig. 9. DoF of the Type I channel in Regime 1 (Bursty A:  $p_d = 0.7, p_c = 0.3, p_{d|c} = 0.5$ , Bursty B:  $p_d = 0.9, p_c = 0.1, p_{d|c} = 0.5$ )

## V. THE UPPER BOUND

We now prove the upper bound of the sum DoF of the bursty MIMO X channel given by Theorem 1. The basic idea of the proof is to extend the state-sequence pairing technique in [7] to the more general bursty MIMO X channel. We do this in three major steps. First we provide extra side information to the two receivers, so that the entropy terms of the rate bounds at the two receivers can be combined for subsequent simplification. We then delicately choose a pair of state sequence realizations that can bring the difference of the combined entropy terms to a computable form. Finally, we obtain the upper bound by resorting to the fact that Gaussian distributions maximize differential entropies and applying the law of large numbers to  $\{S_{ij} \mid i, j \in \{1, 2\}\}$ . Before we start, however, we first present a few simple lemmas and facts that will prove useful in bounding the sum DoF.

### A. Preliminaries

The following lemma extends Lemma 2 of [6] to the bursty MIMO X channel, and can be proved along similar lines. Alternatively a simple proof following the arguments for the broadcast channel [16] is shown below.

*Lemma 1:* The capacity region of the bursty MIMO X channel without feedback depends on distribution of  $(S_{11}, S_{12}, S_{21}, S_{22})$  only through the marginal distributions of  $(S_{11}, S_{12})$  and  $(S_{21}, S_{22})$ .

*Proof:* Let  $P_e \triangleq \{M_{ji} \neq \hat{M}_{ji}, \text{ for some } i, j \in \{1, 2\}\}$ , and  $P_{ej} \triangleq \{M_{ji} \neq \hat{M}_{ji}, \text{ for some } i \in \{1, 2\}\}, j = 1, 2$ . It is easily verified that, for any sequence of coding and decoding schemes,  $P_e$  tends to zero if and only if  $P_{e1}$  and  $P_{e2}$  both tend to zero. But  $P_{e1}$  and  $P_{e2}$  depends only on the marginal distribution of  $Y_1$  and  $Y_2$ , respectively, and hence in turn only on the marginal distributions of  $(S_{11}, S_{12})$  and  $(S_{21}, S_{22})$ , respectively, for any given set of channel matrices (c.f. Fig. 4). This establishes the lemma. ■

A simple and useful consequence of Lemma 1 is that it allows us to reduce channel states [6]. Specifically, for any bursty MIMO X channel with  $p_c \leq p_d$ , it is easily checked that its marginal distributions of  $(S_{11}, S_{12})$  and  $(S_{21}, S_{22})$  are identical to those of the contracted channel shown in Fig. 10 with only five channel states, where a solid line indicates that the link is on and the probability of each state is given in the parentheses, which will be denoted as  $P_A, P_B, \dots, P_E$  hereafter. Therefore, we shall consider this contracted channel, when deriving the upper bound of the sum DoF.

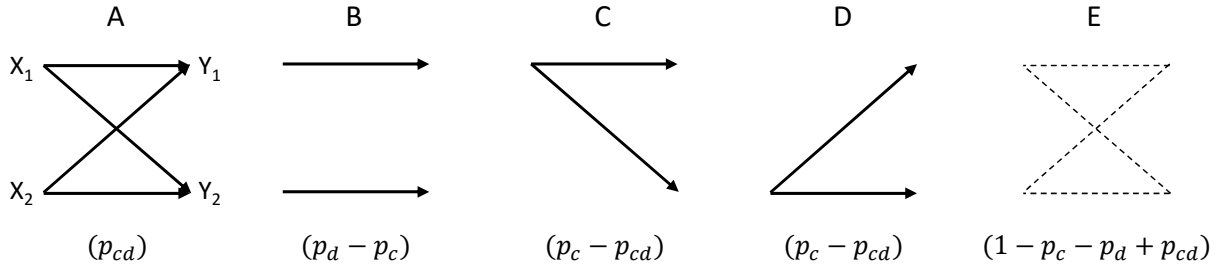


Fig. 10. The contracted channel

Let  $\mathcal{C}(p_c, p_d, p_{cd})$  denote the capacity region of the bursty MIMO X channel without feedback. We have the following lemma as a result of the symmetry of the channel. It follows from this lemma that swapping the values of  $p_c$  and  $p_d$  does not change the sum capacity of the channel, so it suffices to characterize the sum DoF of the channel for  $p_c \leq p_d$ .

*Lemma 2:*  $(R_{11}, R_{12}, R_{21}, R_{22}) \in \mathcal{C}(p_0, p_1, p_2)$  if and only if  $(R_{21}, R_{22}, R_{11}, R_{12}) \in \mathcal{C}(p_1, p_0, p_2)$  for any legitimate set of probabilities  $p_0, p_1$ , and  $p_2$ .

*Proof:* Due to the channel symmetry, the proof is immediate by swapping/relabeling Rx1 and Rx2. ■

Let  $\text{tr}(A)$  denotes the trace of a matrix  $A$ . We collect some well-known or easily-proved facts below, which will facilitate the bounding of the differential entropies. (i) and (ii) are simple; (iii) and (iv) can be proved with techniques to bound conditional differential entropies of jointly Gaussian random vectors, e.g. in [17].

*Fact 1:* Let  $X$  and  $Y$  be two (possibly correlated)  $M$ -dimensional Gaussian random vectors whose correlation matrices are  $R_X$  and  $R_Y$ , respectively, with  $\text{tr}(R_X) \leq kP$  and  $\text{tr}(R_Y) \leq kP$  for some positive constant  $k$ . Let  $Z$  and  $W$  be two independent  $N$ -dimensional Gaussian random vectors, each distributed according to  $\mathcal{N}(0, I_N)$ , independent of  $X$  and  $Y$ . Let  $A, B$  and  $C$  be three  $N \times M$  random matrices with i.i.d. elements drawn from a continuous distribution, independent of  $X, Y, Z$  and  $W$ . Then, with probability 1,

- (i)  $\mathcal{D}\{h(AX + Z)\} \leq \min(M, N)$ ,
- (ii)  $\mathcal{D}\{h(AX + BY + Z)\} \leq \min(2M, N)$ ,
- (iii)  $\mathcal{D}\{h(AX + Z | CX + W)\} \leq \min(M, 2N) - \min(M, N)$ , and
- (iv)  $\mathcal{D}\{h(AX + BY + Z | CY + W)\} \leq \min(M, N)$ .

### B. Providing Extra Side Information to the Receivers

First of all, letting  $S_1 \triangleq (S_{11}, S_{12})$ ,  $S_2 \triangleq (S_{21}, S_{22})$ , and  $S \triangleq (S_1, S_2)$ , the sum rate at Rx1 can be bounded as follows:

$$\begin{aligned}
 & n(R_{11} + R_{12} - \epsilon_n) \\
 & \leq I(M_{11}, M_{12}; Y_1^n, S_1^n) \\
 & \stackrel{(a)}{\leq} I(M_{11}, M_{12}; Y_1^n, S^n, M_{21}) \\
 & \stackrel{(b)}{=} I(M_{11}, M_{12}; Y_1^n | S^n, M_{21}) \\
 & = h(Y_1^n | S^n, M_{21}) - h(Y_1^n | S^n, M_{21}, M_{11}, M_{12}) \\
 & \stackrel{(c)}{=} h(Y_1^n | S^n, M_{21}) - h((S_{12}H_{12}X_2 + Z_1)^n | S^n, M_{12}) \\
 & = \sum_{u^n} \mathcal{P}\{S^n = u^n\} [h(Y_1^n | S^n = u^n, M_{21}) - h((S_{12}H_{12}X_2 + Z_1)^n | S^n = u^n, M_{12})] \tag{2}
 \end{aligned}$$

where side information  $(S_2^n, M_{21})$  is provided to Rx1 in (a), independence between  $(M_{11}, M_{12})$  and  $(S^n, M_{21})$  leads to (b), and (c) follows because  $X_1^n$  becomes deterministic when  $M_{11}$  and  $M_{21}$  are given.

$R_{21} + R_{22}$  is bounded symmetrically, so we also have

$$\begin{aligned} & n(R_{21} + R_{22} - \epsilon_n) \\ & \leq \sum_{v^n} \mathcal{P}\{S^n = v^n\} [h(Y_2^n | S^n = v^n, M_{12}) - h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21})]. \end{aligned} \quad (3)$$

Rearranging the terms, the sum rate is hence bounded from above by

$$\begin{aligned} & n(R_{11} + R_{12} + R_{21} + R_{22} - \epsilon_n) \\ & \leq \left[ \sum_{u^n} \mathcal{P}\{S^n = u^n\} h(Y_1^n | S^n = u^n, M_{21}) - \sum_{v^n} \mathcal{P}\{S^n = v^n\} h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21}) \right] \\ & \quad + \left[ \sum_{v^n} \mathcal{P}\{S^n = v^n\} h(Y_2^n | S^n = v^n, M_{12}) - \sum_{u^n} \mathcal{P}\{S^n = u^n\} h((S_{12}H_{12}X_2 + Z_1)^n | S^n = u^n, M_{12}) \right]. \end{aligned} \quad (4)$$

Next we bound the two differences inside the brackets separately on the contracted channel in Fig. 10.

### C. Pairing State Sequence Realizations

For each difference in (4), we expand the two summations and pair up  $\mathcal{P}\{S^n = u^n\} h(Y_1^n | S^n = u^n, M_{21}) - \mathcal{P}\{S^n = v^n\} h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21})$  with  $u^n$  and  $v^n$  chosen as follows: State C, D and E occur in the same places in  $u^n$  and  $v^n$ , and State A has the same number of occurrences in  $u^n$  and  $v^n$  but has no or minimum overlap, as illustrated in Fig. 11. Note also that in the figure  $\mathcal{I}_u^A$  and  $\mathcal{I}_v^A$  denote the set of time indexes of State A in  $u^n$  and  $v^n$ , respectively. Similar notations are used for other channel states. To keep the expressions in this and next subsection simple, the following notations are also employed:

- (i) Usual set algebra notations:  $|\cdot|$  : cardinality,  $\setminus$  : set minus. E.g.  $|\mathcal{I}_u^A|$ ,  $\mathcal{I}_u^B \setminus \mathcal{I}_v^A$
- (ii)  $\mathcal{I}_u^{AB} \triangleq \mathcal{I}_u^A \cup \mathcal{I}_u^B$  and  $\mathcal{I}_u^{ABC} = \mathcal{I}_u^A \cup \mathcal{I}_u^B \cup \mathcal{I}_u^C$ , and so on.
- (iii)  $X^{\mathcal{I}_u^A} \triangleq \{X[k] \mid k \in \mathcal{I}_u^A\}$ . E.g.  $(H_{11}X_1 + Z_1)^{\mathcal{I}_u^A} = \{(H_{11}X_1[k] + Z_1[k]) \mid k \in \mathcal{I}_u^A\}$

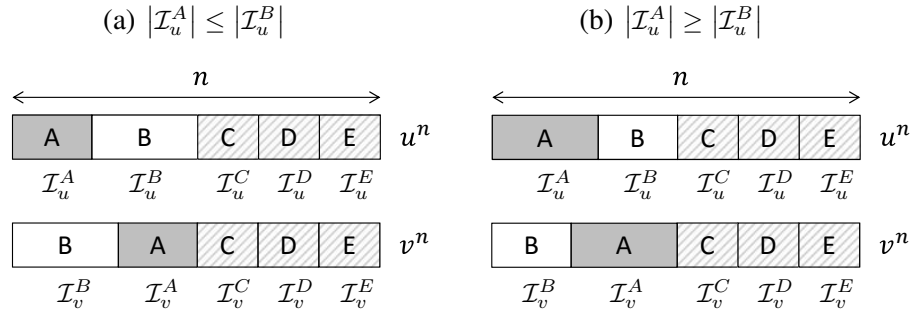


Fig. 11. Pairing of state sequence realizations

With the above pairing strategy, clearly  $\mathcal{P}\{S^n = u^n\} = \mathcal{P}\{S^n = v^n\}$ , and we can bound  $h(Y_1^n | S^n = u^n, M_{21}) - h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21})$  as follows. First note that

$$h(Y_1^n | S^n = u^n, M_{21}) = h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A}, (H_{11}X_1 + Z_1)^{\mathcal{I}_u^{BC}}, (H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}, Z_1^{\mathcal{I}_u^E} \mid M_{21}\right), \text{ and}$$

$$h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21}) = h\left((H_{21}X_1 + Z_2)^{\mathcal{I}_v^{AC}}, Z_2^{\mathcal{I}_v^{BDE}} | M_{21}\right),$$

where  $\{S^n = u^n\}$  and  $\{S^n = v^n\}$  are dropped because they are independent of  $X_1^n$  and  $X_2^n$  when there is no feedback in the system. Letting  $\Omega_1 \triangleq \left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A}, (H_{11}X_1 + Z_1)^{\mathcal{I}_u^{BC}}, (H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}, Z_1^{\mathcal{I}_u^E}\right)$ , and  $\Omega_2 \triangleq \left((H_{21}X_1 + Z_2)^{\mathcal{I}_v^{AC}}, Z_2^{\mathcal{I}_v^{BDE}}\right)$ , we then have

$$\begin{aligned} h(\Omega_1 | M_{21}) &= I(\Omega_1; X_1^n, X_2^n | M_{21}) + h(Z_1^n) \\ &\leq I(\Omega_1, \Omega_2; X_1^n, X_2^n | M_{21}) + h(Z_1^n) \\ &= h(\Omega_1, \Omega_2 | M_{21}) - h(Z_2^n) \\ &= h(\Omega_2 | M_{21}) + h(\Omega_1 | M_{21}, \Omega_2) - h(Z_2^n), \end{aligned}$$

and hence

$$h(Y_1^n | S^n = u^n, M_{21}) - h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21}) \leq h(\Omega_1 | M_{21}, \Omega_2) - h(Z_2^n). \quad (5)$$

For  $|\mathcal{I}_u^A| \leq |\mathcal{I}_u^B|$ , (5) can be further bounded by

$$\begin{aligned} h(Y_1^n | S^n = u^n, M_{21}) - h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21}) \\ \leq h(\Omega_1 | \Omega_2) - h(Z_2^n) \\ \stackrel{(a)}{\leq} h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A}\right) + h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_u^B \setminus \mathcal{I}_u^A}\right) + h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_v^{AC}} | (H_{21}X_1 + Z_2)^{\mathcal{I}_v^{AC}}\right) \\ + h\left((H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}\right), \end{aligned} \quad (6)$$

where (a) follows from the chain rule and the fact that conditioning reduces entropy.

Due to symmetry, we clearly also have

$$\begin{aligned} h(Y_2^n | S^n = v^n, M_{12}) - h((S_{12}H_{12}X_2 + Z_1)^n | S^n = u^n, M_{12}) \\ \leq h\left((H_{21}X_1 + H_{22}X_2 + Z_2)^{\mathcal{I}_v^A}\right) + h\left((H_{22}X_2 + Z_2)^{\mathcal{I}_v^B \setminus \mathcal{I}_v^A}\right) + h\left((H_{22}X_2 + Z_2)^{\mathcal{I}_u^{AC}} | (H_{12}X_2 + Z_1)^{\mathcal{I}_u^{AC}}\right) \\ + h\left((H_{21}X_1 + Z_2)^{\mathcal{I}_v^D}\right). \end{aligned} \quad (7)$$

The case of  $|\mathcal{I}_u^A| \geq |\mathcal{I}_u^B|$  is treated similarly in Appendix A.

#### D. Bounding Differential Entropies

Each of the entropy terms in (6) can in turn be bounded as follows.

$$\begin{aligned} h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A}\right) &\leq \sum_{k \in \mathcal{I}_u^A} h(H_{11}X_1[k] + H_{12}X_2[k] + Z_1[k]) \\ &= |\mathcal{I}_u^A| h(H_{11}X_{1Q_\alpha} + H_{12}X_{2Q_\alpha} + Z_1 | Q_\alpha) \\ &\leq |\mathcal{I}_u^A| h(H_{11}X_{1Q_\alpha} + H_{12}X_{2Q_\alpha} + Z_1) \\ &\leq |\mathcal{I}_u^A| h(H_{11}X_{1G_\alpha} + H_{12}X_{2G_\alpha} + Z_1), \end{aligned} \quad (8)$$

where  $Q_\alpha$  is the time-sharing random variable uniformly distributed on  $\mathcal{I}_u^A$ , and given  $Q_\alpha = k$ ,  $(X_{1Q_\alpha}, X_{2Q_\alpha})$  has the same distribution as  $(X_1[k], X_2[k])$  for any  $k \in \mathcal{I}_u^A$ .  $(X_{1G_\alpha}, X_{2G_\alpha})$  are jointly Gaussian with the same mean and covariance matrix as those of  $(X_{1Q_\alpha}, X_{2Q_\alpha})$ . The other entropy terms in (6) can be similarly bounded for  $P_A \leq P_B$ , and it is straightforward to show

$$\begin{aligned} h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_u^B \setminus \mathcal{I}_v^A}\right) &\leq (|\mathcal{I}_u^B| - |\mathcal{I}_v^A|) h(H_{11}X_{1G_\beta} + Z_1), \\ h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_v^{AC}} \mid (H_{21}X_1 + Z_2)^{\mathcal{I}_v^{AC}}\right) &\leq (|\mathcal{I}_v^A| + |\mathcal{I}_v^C|) h((H_{11}X_{1G_\gamma} + Z_1) \mid (H_{21}X_{1G_\gamma} + Z_2)), \text{ and} \\ h\left((H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}\right) &\leq |\mathcal{I}_u^D| h(H_{12}X_{2G_\delta} + Z_1), \end{aligned} \quad (9)$$

where  $X_{1G_\beta}$ ,  $X_{1G_\gamma}$ , and  $X_{2G_\delta}$  apparently parallel the definitions in (8). Note that we have used the fact that Gaussian distributions maximize (conditional) differential entropies [16, Ch2], [18, Lemma 1].

Now let  $\mathcal{T}_\epsilon^{(n)}$  be the set of typical sequences of  $S^n$ , i.e. those  $u^n$  with  $|\mathcal{I}_u^s|/n - P_s| < \epsilon$  for all  $s \in \{A, B, C, D, E\}$ . Then  $\mathcal{P}\{S^n \notin \mathcal{T}_\epsilon^{(n)}\} = \delta_n \rightarrow 0$  as  $n \rightarrow \infty$ , for any  $\epsilon > 0$ , due to the law of large numbers. As a result, together with Fact 1, it is not hard to verify that, for  $P_A \leq P_B$ , we have

$$\begin{aligned} &\frac{1}{n} \mathcal{D}\{h(Y_1^n \mid S^n, M_{21}) - h((SH_{21}X_1 + Z_2)^n \mid S^n, M_{21})\} \\ &\leq (P_A + \epsilon) \min(2M, N) + (P_B - P_A + 2\epsilon) \min(M, N) + (P_A + P_C + 2\epsilon)[\min(M, 2N) - \min(M, N)] \\ &\quad + (P_D + \epsilon) \min(M, N) + \delta_n K_2 \\ &= P_A \min(2M, N) + (P_A + P_C) \min(M, 2N) + (P_B - 2P_A) \min(M, N) + \epsilon K_1 + \delta_n K_2 \quad (\text{a.s.}), \end{aligned} \quad (10)$$

where  $K_1$  and  $K_2$  are some finite constants. ( $\delta_n K_2$  bounds the contribution of the non-typical sequences.) Since  $\epsilon$  and  $\delta_n$  can both be made arbitrarily small as  $n \rightarrow \infty$ , this implies that

$$\begin{aligned} &\frac{1}{n} \mathcal{D}\{h(Y_1^n \mid S^n, M_{21}) - h((SH_{21}X_1 + Z_2)^n \mid S^n, M_{21})\} \\ &\leq P_A \min(2M, N) + (P_A + P_C) \min(M, 2N) + (P_B - 2P_A) \min(M, N) \\ &= p_{cd} \min(2M, N) + p_c \min(M, 2N) + (p_d - p_c - 2p_{cd}) \min(M, N) \quad (\text{a.s.}). \end{aligned} \quad (11)$$

Finally, recalling the symmetry between (6) and (7), we have proved Theorem 1 for Regime 1  $\left(p_c \leq \frac{p_d}{1+p_{d|c}}\right)$ , which is equivalent to  $P_A \leq P_B$ . The proof of Theorem 1 for Regime 2  $\left(p_c > \frac{p_d}{1+p_{d|c}}\right)$  is obtained by considering  $P_A > P_B$ , following the same lines, and is relegated to Appendix A.

## VI. PROOF OF THEOREM 2 : ACHIEVABILITY SCHEMES

In this section, we show that the upper bound derived in the previous section is in fact tight in Regime 1 and partially tight in Regime 2 (c.f. Fig. 2) with four closely related coding schemes, thereby establishing Theorem 2. Not surprisingly the INBF schemes in [4] that achieve the best-known DoF on the non-bursty MIMO X channel serve as our baseline schemes, and it turns out that they still work well for the Type I channel in Regime 2 even when the channel becomes bursty. However, new beamforming schemes are needed to achieve the sum DoF of the bursty MIMO X channel in Regime 1, where  $p_c$  is usually small. Moreover, for the Type II channel operated



in Regime 2, the INBF schemes alone no longer achieve the DoF, when  $N/M > 1/2$ . Further rate splitting and introducing public messages can offer higher DoF and, in some cases, achieve the DoF of the Type II channel in Regime 2.

Our proof in this section makes use of the simple facts in Fact 2. It is easy to see from Fact 2 that when the channel state  $S$  takes the form of Fig. 6, the achievable DoF can be expressed in the  $\eta$ -triple notation as  $\langle \text{rank}\{H(p_{cd})\}, \text{rank}\{H(p_{\bar{c}d})\}, \text{rank}\{H(p_{c\bar{d}})\} \rangle_{\mathcal{P}_S}$ . To facilitate the proof, it helps to express the upper bound also in the  $\eta$ -triple notation, as shown in Table II.

*Fact 2:* For a point-to-point MIMO Gaussian channel, modeled by  $Y = H(S)X + Z$ , where the channel matrix  $H(S)$  varies with a discrete channel state  $S$  known to the receiver, the following DoF is achievable:  $\sum_s \mathcal{P}_S(s) \text{rank}\{H(s)\}$ . Similarly, for a MIMO Gaussian MAC, modeled by  $Y = H_1(S)X_1 + H_2(S)X_2 + Z$ , where the channel state  $S$  is known to the receiver, if we let  $H(S) = [H_1(S) \ H_2(S)]$  and  $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ , then exactly the same form of sum DoF, i.e.  $\sum_s \mathcal{P}_S(s) \text{rank}\{H(s)\}$ , is also achievable on the MAC.

We now show that the beamforming schemes illustrated in Fig. 12 suffice to achieve the sum DoF upper bound of the Type I bursty MIMO X channel in Regime 1 and part of Regime 2. Note that details are given only for beamforming filters at Tx1 and Rx1 in Fig. 12, Fig. 13, and Fig. 14. It is understood that the beamforming at Tx2 and Rx2 is done symmetrically.

#### A. Type I, Regime 1 : INBF without cross-link messaging

Let us consider the beamforming scheme shown in Fig. 12(a) for the Type I channel operated in Regime 1 with  $1/2 < M/N \leq 1$  first. (The achievability scheme for  $M/N \leq 1/2$  is simpler and can be viewed as a special degenerate case.) Since  $p_c$  is generally small in Regime 1, we essentially have two orthogonal point-to-point channels on the direct-links for a large fraction of time. So a simple and natural coding strategy in this regime is to treat cross-link interference as erasure and use the channel as two orthogonal point-to-point channels on the direct-links. This is the basic idea of Fig. 12(a), where four linear beamforming filters (matrices) are employed.  $\psi_1$  and  $\psi_2$  filters null out the cross-link interference at the receivers, resulting in two effective point-to-point  $M \times (N - M)$  channels, free of interference.  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  filters, on the other hand, exploit the opportunity for data transmission when the cross-links are off, thereby increasing the dimensions of the effective channels to  $M \times M$ .  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  outputs are discarded when the cross-links are present.

Table II  
SUM DOF UPPER BOUND IN THE  $\eta$ -TRIPLE NOTATION

| Type | $\frac{M}{N}$ or $\frac{N}{M}$     | Regime 1                           | Regime 2                            |
|------|------------------------------------|------------------------------------|-------------------------------------|
| I    | $0 < \frac{M}{N} \leq \frac{1}{2}$ | $2\langle M, M, 0 \rangle$         | $2\langle M, M, 0 \rangle$          |
|      | $\frac{1}{2} < \frac{M}{N} \leq 1$ | $2\langle N - M, M, 0 \rangle$     | $2\langle M, N - M, 2M - N \rangle$ |
| II   | $0 < \frac{N}{M} \leq \frac{1}{2}$ | $2\langle N, N, N \rangle$         | $2\langle N, N, N \rangle$          |
|      | $\frac{1}{2} < \frac{N}{M} \leq 1$ | $2\langle M - N, N, M - N \rangle$ | $2\langle N, M - N, N \rangle$      |

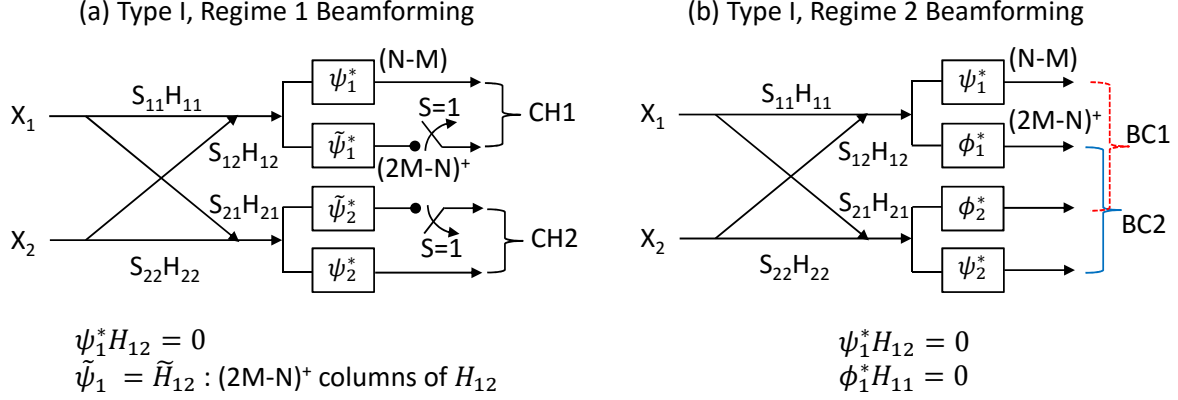


Fig. 12. Beamforming schemes that achieve the DoF upper bound of the Type I channel ( $(x)^+ \triangleq \max(x, 0)$ )

More precisely, we choose  $\psi_1$  to be an  $N \times (N - M)$  matrix, satisfying  $\psi_1^* H_{12} = 0$ . This is certainly feasible since the  $\text{rank}(H_{12}) = M$  (a.s.), and  $\psi_1$  can be made of the  $(N - M)$  basis vectors of the orthogonal complement of the column space of  $H_{12}$ , i.e.  $(\text{Col } H_{12})^\perp$ .  $\tilde{\psi}_1$  is designed so that  $[\psi_1 \tilde{\psi}_1]$  is full rank, and may simply consist of any  $(2M - N)$  columns of  $H_{12}$ . It is easy to verify that  $[\psi_1 \tilde{\psi}_1]^* H_{11}$  is full rank (a.s.), so we effectively have a point-to-point time-varying channel from Tx1 to Rx1, where the rank of the effective channel matrix is  $(N - M)$ ,  $M$  and 0 (a.s.) in the  $p_{cd}$ ,  $p_{\bar{c}d}$  and  $p_{c\bar{d}}$  states (Fig. 6), respectively. Therefore, by Fact 2, a DoF of  $\langle N - M, M, 0 \rangle$  is achievable on the link from Tx1 to Rx1. The same DoF is clearly achievable on the Tx2 to Rx2 link due to symmetry. Therefore, we can achieve a total DoF of  $2\langle N - M, M, 0 \rangle$ , which is exactly the respective upper bound in Table II! Note that this scheme maximizes the achievable DoF in the  $p_{\bar{c}d}$  state, while sacrificing some DoF of the  $p_{cd}$  state, which is reasonable because  $p_{cd} \leq p_d - p_c \leq p_{\bar{c}d}$  in Regime 1.

In the case of  $M/N \leq 1/2$ , the  $\psi_1$  and  $\psi_2$  filters can eliminate the interference coming from the other transmitter when the cross-links are present, while preserving the full  $M$  dimensions of the direct-link signals. So the  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  filters are unnecessary, and the rank of each effective point-to-point channel is  $M$  whenever the direct-link is present, even if we reduce the size of the  $\psi_1$  and  $\psi_2$  matrices to  $N \times M$ . It follows that the achievable DoF is  $2\langle M, M, 0 \rangle$  in this case, which matches the upper bound as well. This simplified beamforming scheme is illustrated in Fig. 13(a).

### B. Type I, Regime 2 : INBF with cross-link messaging

For the Type I channel operated in Regime 2, we again consider the  $1/2 < M/N \leq 2/3$  case first. Note that in Regime 2,  $p_{cd}$  and  $p_c$  generally become larger, so the cross-links are present for an increasingly larger fraction of time. Intuitively, to maximize the sum DoF in this case, the paradigm should be shifted from avoiding the cross-links to exploiting the cross-links. One way to accomplish this is cross-link messaging with the beamforming scheme depicted in Fig. 12(b). Note that this is exactly the same INBF scheme for the non-bursty MIMO X channel as proposed in [4].

To determine the DoF achievable by this scheme when the channel becomes bursty, recall that this scheme decomposes the channel into two orthogonal broadcast channels, with  $\psi_1^* H_{12} = \psi_2^* H_{21} = 0$  and  $\phi_1^* H_{11} = \phi_2^* H_{22} = 0$ . Now consider the effective broadcast channel, BC1, from Tx1 to  $\psi_1$  and  $\phi_2$ . Since the effective channel matrix  $\begin{bmatrix} \psi_1^* H_{11} \\ \phi_2^* H_{21} \end{bmatrix}$  is full-rank (a.s.), with a simple channel inversion at Tx1 we have effectively two orthogonal point-to-point channels. It then follows from Fact 2 that sum DoF of  $(N-M)p_d + (2M-N)p_c = \langle M, N-M, 2M-N \rangle_{\mathcal{P}_S}$  can be achieved on BC1. Due to symmetry, the same DoF is also achievable on the other effective broadcast channel (BC2), so the total achievable DoF clearly coincides with the upper bound in Table II, i.e.  $2\langle M, N-M, 2M-N \rangle$ .

When  $M/N \leq 1/2$ , the  $N-M$  dimensions of the  $\psi_1$  filter output can capture all of the  $M$  dimensions of the signal coming from Tx1, and so can the  $\psi_2$  filter. In this case, it does not make sense to split the transmitter dimensions and send messages on the cross-links, since  $p_c \leq p_d$ . So the  $\phi_1$  and  $\phi_2$  filters can be removed, and by sending messages on the direct-links only, we can clearly achieve a DoF of  $2Mp_d = 2\langle M, M, 0 \rangle_{\mathcal{P}_S}$ , which also matches the upper bound. Moreover, the  $\psi_1$  and  $\psi_2$  matrices can be reduced to  $N \times M$  matrices as shown in Fig. 13(a), without affecting the achievable DoF.

It is interesting to note that the two beamforming schemes in Fig. 12 originating from two different paradigms degenerate into the same scheme in Fig. 13(a) when  $M/N \leq 1/2$ , because in this case both paradigms reduce to maximizing the utility of the direct links. This may explain why, when  $M/N \leq 1/2$ , the sum DoF of the channel takes the same form in both regimes, and depends on  $p_d$  only, i.e.  $2\langle M, M, 0 \rangle = 2Mp_d$ .

In summary, with the proof in this and the previous subsection, we have shown that the sum DoF upper bound is indeed tight for the Type I channel in the solid regions of Fig. 2(a), and thereby proving the half of Theorem 2. Next, let us move on to the Type II channel. Since the transmitters have more antennas than the receivers on the Type II channel, a natural coding attempt is to employ beamforming filters on the transmitter side reciprocal to those on the receiver side of the Type I channel. Although this attempt is headed roughly in the right direction, it is not completely correct and optimal for the Type II channel. As we have seen in Section IV, the Type II bursty

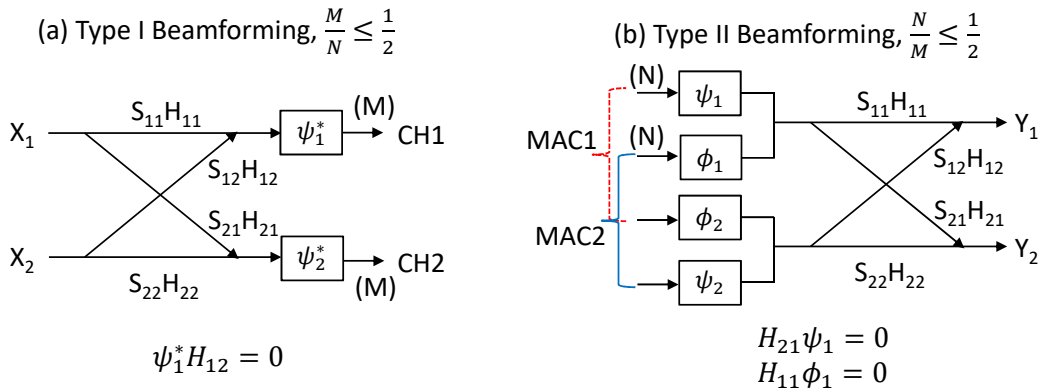


Fig. 13. Simplified beamforming schemes when  $\frac{M}{N} \leq \frac{1}{2}$  or  $\frac{N}{M} \leq \frac{1}{2}$

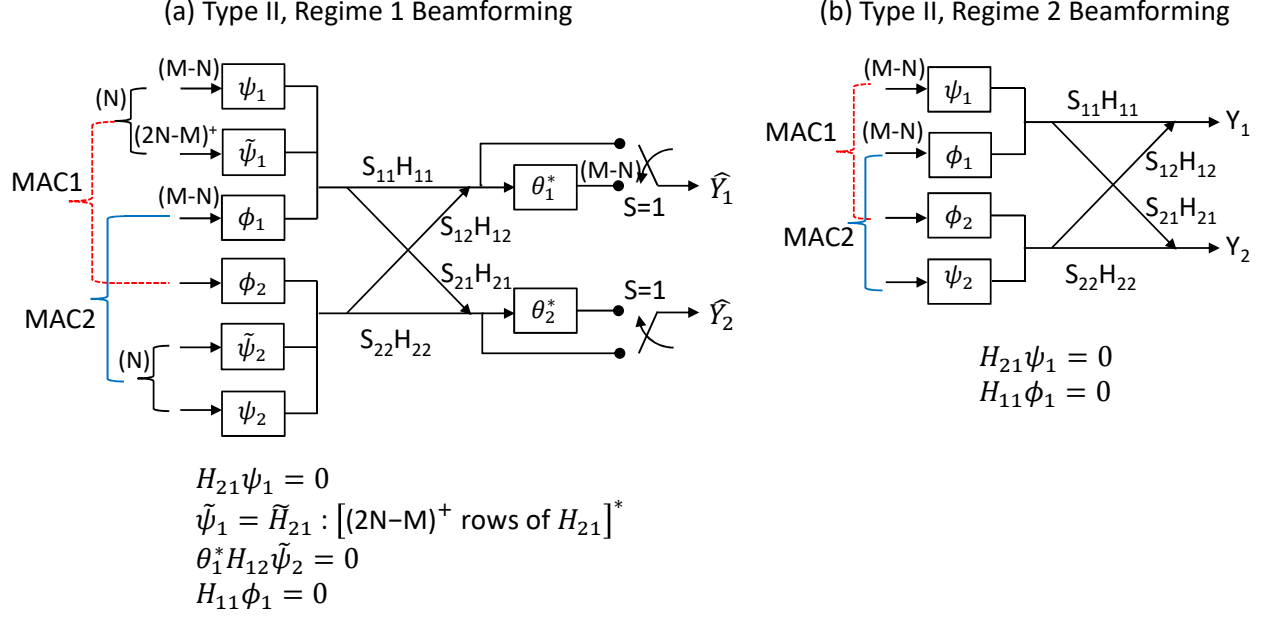


Fig. 14. Beamforming schemes that, optionally together with public messages, achieve the DoF upper bound of the Type II channel

channel is not strictly reciprocal to the Type I channel, and neither is its beamforming scheme, due to a fundamental difference between bursty BC and MAC without feedback, i.e. cross-links play a more important role on the MAC and Type II channel. Therefore, some corrections to the reciprocal scheme are needed, and we will show that with the modifications shown in Fig. 14, and optionally introducing public messages, we can achieve the sum DoF of the bursty Type II channel.

### C. Type II, Regime 1 : INBF with cross-link messaging

Let us start with Regime 1 and  $N/M > 1/2$ . Comparing the beamforming scheme in Fig. 14(a) for the Type II channel to that in Fig. 12 for the Type I channel, we note that they are reciprocal, except for the  $\phi_1$  and  $\phi_2$  filters in Fig. 14(a). Specifically  $\psi_1$  is an  $M \times (M-N)$  matrix satisfying  $H_{21}\psi_1 = 0$ , and  $\tilde{\psi}_1$  of dimensions  $M \times (2N-M)$  is chosen to make  $\begin{bmatrix} \psi_1 & \tilde{\psi}_1 \end{bmatrix}$  full-rank (a.s.)  $\psi_2$  and  $\tilde{\psi}_2$  are designed symmetrically at Tx2. In other words, the  $\psi_1$  and  $\psi_2$  filters are again designed to null out cross-link interference and effect two point-to-point channels on the direct-links, free of interference. The  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  filters enlarge the dimensions of the effective point-to-point links from  $M-N$  to  $N$  when the cross-link are off. However, when the cross-links are on, we need  $\theta_1$  and  $\theta_2$  filters to null out the interference from  $\phi_2$  and  $\phi_1$ , respectively. This can be done with two  $N \times (M-N)$  matrices satisfying  $\theta_1^* H_{12} \phi_2 = \theta_2^* H_{21} \phi_1 = 0$ . This extra work is needed on the Type II channel because only the receivers know the state of the channel. Except for this slight technical difference,  $\psi_1, \tilde{\psi}_1, \psi_2$  and  $\tilde{\psi}_2$  play exactly the same role as in the Type I channel, and present  $\langle M-N, N, 0 \rangle$  DoF on each direct-link.  $(\theta_1^* H_{11} \begin{bmatrix} \psi_1 \\ \tilde{\psi}_1 \end{bmatrix})$  is easily verified to be full-rank almost surely.)

However, more can be done on the Type II channel. Since a Type II channel can be decomposed into two MACs and there are unused dimensions at each transmitter, we incorporate the  $\phi_1$  and  $\phi_2$  filters to take advantage of the extra opportunity of communication offered by the cross-links when the direct-links are off. With  $H_{22}\phi_2 = 0$ , the  $\phi_2$  filter is transparent to Rx2, but it can offer  $(M - N)$  dimensions for data transmission from Tx2 to Rx1 in the  $p_{c\bar{d}}$  state (Fig. 6), due to Fact 2. As such, it is easy to see that the achievable DoF at Rx1 is increased from  $\langle M - N, N, 0 \rangle$  to  $\langle M - N, N, M - N \rangle$ . By symmetry, the same DoF is clearly also achievable at Rx2, so the achievable total DoF is  $2\langle M - N, N, M - N \rangle$ , which is exactly the upper bound of the Type II channel in Regime 1 when  $N/M > 1/2$ ! The reader may have noticed a subtle issue here: Is  $\theta_1^* H_{12} \phi_2$  full rank (a.s.)? Although  $\theta_1$  and  $H_{12}$  are correlated, the answer is still yes, and can be seen in Appendix B. Similarly,  $\theta_2^* H_{21} \phi_1$  is also full rank (a.s.)

When  $N/M \leq 1/2$ , similarly the  $\psi_1$  and  $\psi_2$  filters alone can deliver  $N$  dimensions of signals to Rx1 and Rx2, respectively, whenever the direct-links are on. So the  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  filters again become redundant and can be removed. Moreover, the dimensions of the  $\psi_1, \psi_2, \phi_1$  and  $\phi_2$  filters can all be reduced to  $M \times N$ , and still achieve the DoF upper bound  $2\langle N, N, N \rangle$ . This simplified scheme is shown in Fig. 13(b).

#### D. Type II, Regime 2 : INBF with cross-link messaging, overlaid with public messages (HKIA)

1) *Sub-optimality of INBF*: Finally we come to the more intriguing case of the Type II channel in Regime 2, where INBF may not achieve the DoF of the channel. To see this, let us consider the case of  $1/2 < N/M \leq 2/3$  first. Following the philosophy of the paradigm shift on the Type I channel from avoiding the cross-links to exploiting them, we remove the  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$  filters and use the  $(2N - M)$  dimensions at each transmitter to beamform (with  $\phi_1$  and  $\phi_2$ ) to the cross-link receiver. ( $\theta_1$  and  $\theta_2$  filters should therefore also be removed.) Moreover, the dimensions of the  $\phi_1$  and  $\phi_2$  filters can be dialed up from  $M \times (2N - M)$  to  $M \times (M - N)$ . This leads us to the beamforming scheme in Fig. 14(b). With Fact 2, it is straightforward to verify that this scheme can deliver  $2\langle N, M - N, M - N \rangle$  DoF, which is already greater than the reciprocal of the DoF of the corresponding Type I channel, i.e.  $\langle N, M - N, 2N - M \rangle$ . The difference is apparently due to the increased dimensions of the  $\phi_1$  and  $\phi_2$  filters, and is another manifestation of the bursty BC-MAC difference.

At this point, we have used up the available dimensions for beamforming to orthogonalize the channel. However, the DoF upper bound is not attained, and an examination of the signal dimensions at the receivers, as was done in Section IV, quickly reveals that we fall short of the potential of the channel: The received signal dimensions at each receiver are  $N, (M - N)$  and  $(M - N)$  in the  $p_{cd}, p_{\bar{c}d}$  and  $p_{c\bar{d}}$  states, respectively (c.f. Fig. 6), but there are  $N$  Rx antennas. So the  $p_{\bar{c}d}$  and  $p_{c\bar{d}}$  states are underutilized by the INBF scheme, since  $(M - N) < N \leq M$  when  $1/2 < N/M \leq 2/3$ .

2) *DoF-achieving HKIA scheme*: One way to make full use of the under-utilized states is the HKIA scheme which employs rate splitting, similar to the HK strategy in [9]. Specifically, we introduce public messages,  $M_{01}$  and  $M_{02}$ , and superimpose them on the INBF signals in Fig. 14(b) which carry the private messages. ( $M_{01}$  and  $M_{02}$  are the messages from Tx1 and Tx2, respectively, to both receivers.) The transmit power is split evenly between

public and private messages. More precisely, we have the following random coding scheme, generalized from the one used in Section IV-A. (For completeness, we repeat some of the earlier definitions and arguments below.)

$$X_1^n(M_{01}, M_{11}, M_{21}) = U_1^n(M_{01}) + [\psi_1 \ \phi_1] \begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}, \quad (12)$$

where  $U_1 \sim \mathcal{N}(0, \frac{P}{2M} I_M)$ ,  $C_1 \& D_1 \sim \mathcal{N}(0, \frac{P}{4(M-N)} I_{M-N})$ , and  $[\psi_1 \ \phi_1] \begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}$  is a shorthand notation for multiplying each of the  $n$  components of  $\begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}$  by  $[\psi_1 \ \phi_1]$ .  $X_2^n(M_{02}, M_{12}, M_{22})$  is encoded symmetrically:

$$X_2^n(M_{02}, M_{12}, M_{22}) = U_2^n(M_{02}) + [\psi_2 \ \phi_2] \begin{bmatrix} D_2^n(M_{22}) \\ C_2^n(M_{12}) \end{bmatrix}, \quad (13)$$

where  $U_2, C_2$ , and  $D_2$  have the same distributions as  $U_1, C_1$ , and  $D_1$ . Successive decoding is employed at each receiver, where the public messages are decoded and removed first, and then the private messages are decoded. Note that although this scheme is motivated from  $1/2 < N/M \leq 2/3$ , there is nothing preventing us from using it for  $2/3 < N/M \leq 1$ . The achievable sum DoF of this HKIA scheme is given in Proposition 1. To facilitate the discussion hereafter, we shall use the terms *private DoF* ( $\eta_{\text{priv}}$ ) and *public DoF* ( $\eta_{\text{pub}}$ ) to refer to the sum DoF achievable on the private messages and public messages, respectively.

*Proposition 1:* The following private and public DoF can be achieved by the HKIA scheme in Section VI-D2:

- (i)  $\eta_{\text{priv}} = 2 \langle N, M - N, M - N \rangle$  and  $\eta_{\text{pub}} = 2 \langle 0, 0, 2N - M \rangle$ ; when  $1/2 < N/M \leq 2/3$ .
- (ii)  $\eta_{\text{priv}} = 2(M - N) \langle 2, 1, 1 \rangle$  and  $\eta_{\text{pub}} = 2 \langle 3N - 2M, 0, 2N - M \rangle$ , when  $2/3 < N/M \leq 1$  and  $p_c \leq \frac{p_d}{1 + \alpha p_{d|c}}$ , where  $\alpha = \frac{3N - 2M}{2N - M}$ . In both cases, the HKIA scheme achieves a total DoF of  $2 \langle N, M - N, N \rangle$ .

Note that non-zero public DoF is achieved whenever  $p_{c\bar{d}} > 0$ , and the total achievable DoF matches the upper bound given in Table II.

3) *Proof of Proposition 1 for  $1/2 < N/M \leq 2/3$ :* First of all, public messages are removed when we decode the private messages, so the achievable private DoF clearly remains the same as in Section VI-D1, which readily proves that  $\eta_{\text{priv}} = 2 \langle N, M - N, M - N \rangle$  is achievable. So the key task is to evaluate the achievable  $\eta_{\text{pub}}$ . To do so, let  $\eta_{01}$  and  $\eta_{02}$  denote the achievable DoF on  $M_{01}$  and  $M_{02}$ , respectively. Since the public messages are decoded at both receivers (or effectively two MACs) in the presence of the private messages,  $(\eta_{01}, \eta_{02})$  pair satisfying the following set of inequalities are achievable (i.e. the intersection of the DoF regions of the two MACs):

$$\begin{aligned} \eta_{01} &\leq \mathcal{D}\{\min[I(U_1; Y_1 \mid S_1, U_2), I(U_1; Y_2 \mid S_2, U_2)]\}, \\ \eta_{02} &\leq \mathcal{D}\{\min[I(U_2; Y_1 \mid S_1, U_1), I(U_2; Y_2 \mid S_2, U_1)]\}, \text{ and} \\ \eta_{01} + \eta_{02} &\leq \mathcal{D}\{\min[I(U_1, U_2; Y_1 \mid S_1), I(U_1, U_2; Y_2 \mid S_2)]\}, \end{aligned} \quad (14)$$

Each of the mutual informations in (14) can be expressed as the difference of two conditional differential entropies, e.g.  $I(U_1; Y_1 \mid S, U_2) = h(Y_1 \mid S, U_2) - h(Y_1 \mid S, U_1, U_2)$ . So we first compute  $\mathcal{D}\{h(Y_1 \mid S_1)\}$ ,  $\mathcal{D}\{h(Y_1 \mid$

$S_1, U_1\}$ ,  $\mathcal{D}\{h(Y_1 | S_1, U_2)\}$ , and  $\mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\}$ . The  $\eta$ -triple notation comes particularly handy in this analysis. Following the steps given in Appendix C, it is not hard to show that

$$\begin{aligned}\mathcal{D}\{h(Y_1 | S_1)\} &= \langle N, N, N \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_1)\} &= \langle N, M - N, N \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_2)\} &= \langle N, N, M - N \rangle, \text{ and} \\ \mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} &= \langle N, M - N, M - N \rangle.\end{aligned}\tag{15}$$

By symmetry, we clearly also have

$$\begin{aligned}\mathcal{D}\{h(Y_2 | S_2)\} &= \langle N, N, N \rangle, \\ \mathcal{D}\{h(Y_2 | S_2, U_2)\} &= \langle N, M - N, N \rangle, \\ \mathcal{D}\{h(Y_2 | S_2, U_1)\} &= \langle N, N, M - N \rangle, \text{ and} \\ \mathcal{D}\{h(Y_2 | S_2, U_1, U_2)\} &= \langle N, M - N, M - N \rangle.\end{aligned}\tag{16}$$

With (15) and (16), it is then straightforward to show that (14) reduces to

$$\begin{aligned}\eta_{01}, \eta_{02} &\leq \langle 0, 0, 2N - M \rangle, \text{ and} \\ \eta_{01} + \eta_{02} &\leq \langle 0, 2N - M, 2N - M \rangle.\end{aligned}\tag{17}$$

From this we conclude that  $\eta_{\text{pub}} = \eta_{01} + \eta_{02}$  satisfying the following inequality is achievable:

$$\eta_{\text{pub}} \leq \min \left( 2\langle 0, 0, 2N - M \rangle, \langle 0, 2N - M, 2N - M \rangle \right) = 2\langle 0, 0, 2N - M \rangle,\tag{18}$$

where the last equality holds because  $p_c \leq p_d$ . This establishes Proposition 1 for  $1/2 < M/N \leq 2/3$ .

4) *Proof of Proposition 1 for  $2/3 < N/M \leq 1$ :* Again, it is easy to see that the following private DoF is achievable when  $2/3 < N/M \leq 1$ :

$$\eta_{\text{priv}} = 2(M - N) \langle 2, 1, 1 \rangle\tag{19}$$

As for the public DoF, it is not hard to verify that (15) and (16) remain valid, except for the following modification:

$$\mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} = \mathcal{D}\{h(Y_2 | S_2, U_1, U_2)\} = \langle 2M - 2N, M - N, M - N \rangle.\tag{20}$$

As a result, (14) evaluates to

$$\begin{aligned}\eta_{01}, \eta_{02} &\leq \langle 3N - 2M, 0, 2N - M \rangle, \text{ and} \\ \eta_{01} + \eta_{02} &\leq \langle 3N - 2M, 2N - M, 2N - M \rangle,\end{aligned}\tag{21}$$

which implies that the following public DoF is achievable:

$$\eta_{\text{pub}} = \eta_{01} + \eta_{02} \leq \min \left( 2\langle 3N - 2M, 0, 2N - M \rangle, \langle 3N - 2M, 2N - M, 2N - M \rangle \right).\tag{22}$$

Note that  $2\langle 3N - 2M, 0, 2N - M \rangle_{\mathcal{P}_S} \leq \langle 3N - 2M, 2N - M, 2N - M \rangle_{\mathcal{P}_S}$  when  $(3N - 2M)p_{cd} \leq (2N - M)(p_d - p_c)$ , or equivalently  $p_c \leq \frac{p_d}{1 + \alpha p_{d|c}}$ , where  $\alpha = \frac{3N - 2M}{2N - M}$ . Therefore,  $\eta_{\text{pub}} \leq 2\langle 3N - 2M, 0, 2N - M \rangle$  is achievable when  $p_c \leq \frac{p_d}{1 + \alpha p_{d|c}}$ , which concludes the proof of Proposition 1.

5) *Simplified scheme for  $N/M \leq 1/2$* : When  $N/M \leq 1/2$ , note that the beamforming scheme in Fig. 14(b) has completely used up the  $N$  dimensions of the receivers in each of  $p_{cd}$ ,  $p_{\bar{c}d}$ , and  $p_{c\bar{d}}$  states. Overlaying the private messages with public messages in this case is hence unfruitful and unnecessary. The private messages alone can deliver a sum DoF of  $2\langle N, N, N \rangle$ , which already matches the upper bound. Moreover, the dimensions of the  $\psi_1, \phi_1, \psi_2$  and  $\phi_2$  filters can all be reduced to  $M \times N$ , resulting in the simplified scheme in Fig. 13(b).

Once again, when  $N/M \leq 1/2$ , the two beamforming schemes in Fig. 14 for the Type II channel both reduce to the same scheme in Fig. 13(b), and the sum DoF of the channel has the same form, i.e.  $2\langle N, N, N \rangle_{\mathcal{P}_S} = 2Np_d + 2Np_{\bar{c}d}$ , in Regime 1 and Regime 2. Unlike the Type I channel, however, the sum DoF in this case varies with  $p_c$  and  $p_{cd}$ , due to the fundamental difference between the bursty BC and MAC discussed earlier.

To recapitulate this section, we have shown four closely related coding schemes that attain the sum DoF upper bound of the bursty MIMO X channel in the solid regions of Fig. 2, and hence have proved Theorem 2.

## VII. PROOF OF THEOREM 3 : BLENDING HKIA SCHEMES AND FINE-TUNING POWER ALLOCATION

We now turn to the proof of Theorem 3, which furnishes a lower bound of the sum DoF of the bursty MIMO X channel where the sum DoF remains unknown, i.e. in the dotted regions of Fig. 2. This lower bound is obtained by further extending the HKIA scheme shown in Section VI-D, and it surpasses the sum DoF achievable by both the IA scheme and the HKIA scheme in Section VI-D. To differentiate HKIA schemes with different INBF schemes on a bursty MIMO X channel in this section, we shall use  $[N_{D_1}, N_{C_1}, N_{C_2}, N_{D_2}]_{\text{II}}$  to refer to an HKIA scheme on the Type II channel where the  $D_1, C_1, C_2$ , and  $D_2$  random vectors take  $N_{D_1}, N_{C_1}, N_{C_2}$ , and  $N_{D_2}$  dimensions, respectively, c.f. (12) and (13). Similarly, the  $[N_{D_1}, N_{C_1}, N_{C_2}, N_{D_2}]_{\text{I}}$  notation shall also be used to distinguish the HKIA schemes on the Type I channel.

### A. Motivation

To motivate the general scheme, let us first consider a concrete and simple example, namely the  $6 \times 5$  bursty X channel with  $\frac{p_d}{1 + \alpha p_{d|c}} < p_c \leq p_d$ , where  $\alpha = \frac{3N - 2M}{2N - M} = \frac{3}{4}$ . For this channel, we know from (19) and (22) that the HKIA scheme ( $[1, 1, 1, 1]_{\text{II}}$ ) in Section VI-D achieves a DoF of  $2(M - N)\langle 2, 1, 1 \rangle + \langle 3N - 2M, 2N - M, 2N - M \rangle = \langle 7, 6, 6 \rangle_{\mathcal{P}_S}$ . However, a moment of reflection reveals that this clearly cannot be optimal. For example, when  $p_c, p_d$  and  $p_{cd}$  are all close to 1, the achieved sum DoF is  $\langle 7, 6, 6 \rangle_{\mathcal{P}_S} \simeq 7$ . But in this case we almost have a non-bursty  $6 \times 5$  X channel, and by using the usual INBF (IA) scheme (Fig. 16(b)) for the non-bursty channel alone we can easily achieve  $\langle 8, 4, 4 \rangle_{\mathcal{P}_S} \simeq 8$  DoF, which should also be close to being optimal. So intuitively to enhance the lower bound we would consider an alternative HKIA scheme ( $[2, 2, 2, 2]_{\text{II}}$ ) with the INBF scheme in Fig. 16(b). This scheme differs from the previous HKIA scheme only in private message signaling, with  $D_1, C_1, C_2, D_2 \sim \mathcal{N}(0, \frac{P}{4M/3} I_{M/3})$  in (12) and (13). Following the same steps in Section VI-D, it is easy to show that this HKIA achieves a DoF of



$\langle 8, 4, 8 \rangle_{\mathcal{P}_S}$ , which is strictly greater than  $\langle 8, 4, 4 \rangle_{\mathcal{P}_S}$  whenever  $p_{cd} > 0$ . These three lower bounds are illustrated in Fig. 15.

Rather than just taking the maximum of  $\langle 7, 6, 6 \rangle_{\mathcal{P}_S}$  and  $\langle 8, 4, 8 \rangle_{\mathcal{P}_S}$  to be the new lower bound, we consider a simple way to *blend* these two HKIA schemes that yields an even better bound. Specifically, let each of  $D_1, C_1, C_2$  and  $D_2$  now takes  $\mathcal{N}\left(0, \begin{bmatrix} \frac{1}{8}P & 0 \\ 0 & \frac{1}{8}P^a \end{bmatrix}\right)$  distribution, where  $a \in [0, 1]$  is a power tuning factor. Note that, with  $a = 1$ , it is exactly the  $[2, 2, 2, 2]_{\text{II}}$  HKIA scheme, while it yields the  $[1, 1, 1, 1]_{\text{II}}$  HKIA lower bound when  $a = 0$ . Clearly, with this scheme, tuning up ‘ $a$ ’ would increase the private DoF but decrease the public DoF, and vice versa. So by optimizing ‘ $a$ ’ for any given  $p_c, p_d$  and  $p_{cd}$ , we can find a good trade-off between private and public DoF and obtain a lower bound that is guaranteed to be no less than  $\max(\langle 7, 6, 6 \rangle_{\mathcal{P}_S}, \langle 8, 4, 8 \rangle_{\mathcal{P}_S})$ . As shown in Appendix D, this HKIA scheme achieves a DoF of  $6\left(\langle 1, 1, 1 \rangle_{\mathcal{P}_S} + \langle 1, 0, 0 \rangle_{\mathcal{P}_S} \frac{\langle 1, -1, 1 \rangle_{\mathcal{P}_S}}{\langle 3, -2, 2 \rangle_{\mathcal{P}_S}}\right)$  and is illustrated as ‘Proposed’ in Fig. 15. Compared to the  $[1, 1, 1, 1]_{\text{II}}$  HKIA scheme, this scheme performs markedly better because it trades off between the DoF achievable on the public and private messages by exploring two beamforming options, and fine-tuning the power allocation.

*Remark 1:* In the spirit of trading off between public and private DoF, one may want to introduce another power tuning factor ‘ $b$ ’ and consider  $\mathcal{N}\left(0, \begin{bmatrix} \frac{1}{8}P^b & 0 \\ 0 & \frac{1}{8}P^a \end{bmatrix}\right)$  for  $D_1, C_1, C_2$  and  $D_2$ . This, however, is unnecessary, because the optimal value of ‘ $b$ ’ is in fact unity, as shown in Appendix D.

*Remark 2:* Note also that we fine-tune the power of the private messages, but not the public messages, because reducing the power of the public messages does not help the private DoF for our successive decoding scheme.

*Remark 3:* For ease of presentation, we shall reverse the usual order and treat the Type II channel first. The HKIA scheme for the Type I channel will be shown in Section VII-C.

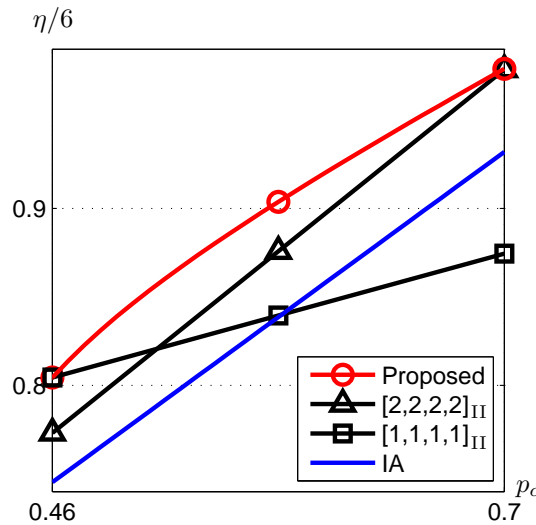


Fig. 15. Three HKIA schemes for the 6x5 bursty channel with  $p_d = 0.7, p_{d|c} = 0.9, \frac{p_d}{1+\alpha p_{d|c}} < p_c \leq p_d$

### B. HKIA Lower Bound for the Type II Channel

For  $\frac{2}{3} < \frac{N}{M} \leq 1$  and  $\frac{p_d}{1+\alpha p_{d|c}} < p_c \leq p_d$ , when  $M$  is a multiple of 3, the general HKIA scheme blends the  $[(M-N), (M-N), (M-N), (M-N)]_{\text{II}}$  and the  $[\frac{M}{3}, \frac{M}{3}, \frac{M}{3}, \frac{M}{3}]_{\text{II}}$  INBF schemes, c.f. Fig. 14(b) and Fig. 16(b). The random coding scheme takes the same form as the one in Section VI-D2, with only the following changes to the private message encoding: We employ the INBF in Fig. 16(b), where the input vector to each beamforming matrix is partitioned into two groups. The first  $(M-N)$  components are allocated with average power of  $P^b$  and are beamformed in such a way that they cause no interference to the unintended receiver. The rest  $(N-2M/3)$  components are transmitted with  $P^a$  average power, and their beamforming nulls out interference only partially at the unintended receiver. More specifically,  $D_1, C_1, C_2$  and  $D_2$  in (12) and (13) now each takes the following distribution:

$$D_1, C_1, C_2, D_2 \sim \mathcal{N} \left( 0, \begin{bmatrix} \frac{1}{4(M-N)} P^b I_{M-N} & 0 \\ 0 & \frac{1}{4(N-2M/3)} P^a I_{N-2M/3} \end{bmatrix} \right) \quad (\text{i.i.d.}), \quad (23)$$

where  $a, b \in [0, 1]$ . This HKIA scheme is extended in Appendix D for arbitrary  $M = 3k + q$ , where  $k$  is an integer and  $q \in \{0, 1, 2\}$ . As a result we have the following proposition, which establishes  $\eta_{\text{lb},2}$  of Theorem 3.

*Proposition 2:* For the HKIA scheme in Section VII-B and Appendix D, the optimal value of ‘ $b$ ’ is 1. The optimal value of ‘ $a$ ’ and the achievable sum DoF are as follows ( $\beta_2 = \frac{q}{k+q}$ ):

$$a_{\text{opt}} = \begin{cases} \frac{\langle 3N - 2M, M - 2N, 2N - M \rangle}{(k - M + N) \langle 3, -2, 2 \rangle}, & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_2 p_{d|c}}, \\ 1, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_2 p_{d|c}}. \end{cases}$$

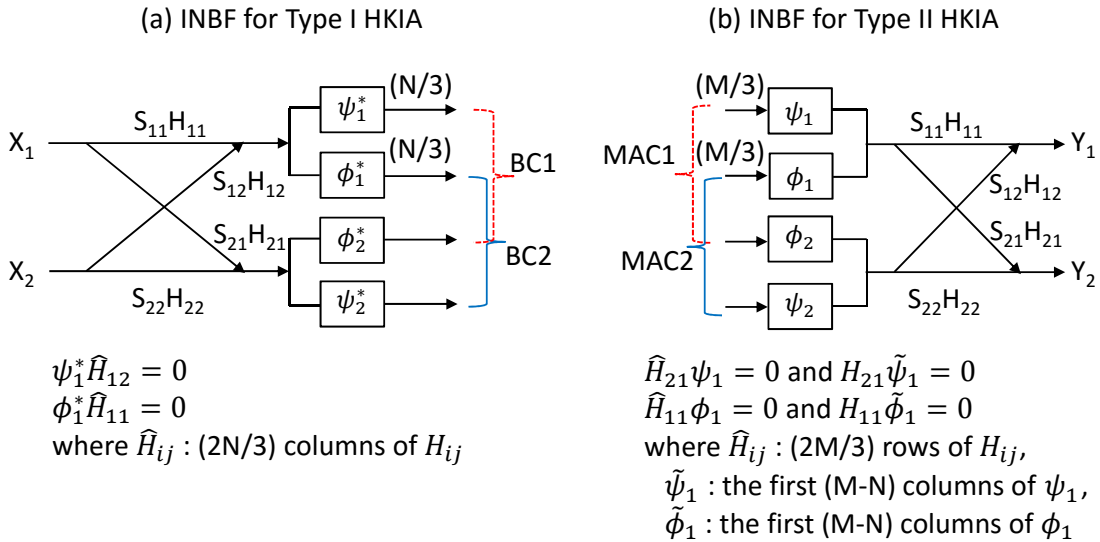


Fig. 16. INBF schemes for the HKIA DoF lower bounds when  $\max(M, N)$  is a multiple of 3

$$\eta_{\text{total}} \leq \begin{cases} M \left( \langle 1, 1, 1 \rangle + \langle 1, 0, 0 \rangle \frac{\langle 1, -1, 1 \rangle}{\langle 3, -2, 2 \rangle} \right), & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_2 p_{d|c}}, \\ M \langle 1, 1, 1 \rangle + k \langle 1, 0, 0 \rangle, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_2 p_{d|c}}. \end{cases}$$

Though somewhat complicated by the more sophisticated INBF and power optimization, the proof of Proposition 2 follows the same lines as in Section VI-D3, and is relegated to Appendix D. The key effort is to re-evaluate the entropies and DoF based on the new distributions and beamforming and to optimize ‘ $a$ ’ and ‘ $b$ ’.

*Remark 4:* It is reasonable that the optimal value of ‘ $b$ ’ is unity. To see this, recall that the first  $(M - N)$  components of  $D_1, C_1, C_2$  and  $D_2$  are allocated with average power  $P^b$  and they cause no interference to the unintended receiver, but not so for the rest  $(N - 2M/3)$  components. In other words, the first  $(M - N)$  components are better beamformed, and it is more effective to send messages on them. This explains why ‘ $b$ ’ is maximized and assumes a larger value than ‘ $a$ ’.

### C. HKIA Lower Bound for the Type I Channel

1) *An Example:* To derive a DoF lower bound for the Type I channel with  $\frac{2}{3} < \frac{M}{N} \leq 1$  and  $\frac{p_d}{1+p_{d|c}} < p_c \leq p_d$ , a natural attempt is to duplicate the schemes on the Type II channel in the previous subsection by constructing HKIA schemes with reciprocal beamforming. Although this strategy again yields a better lower bound for the Type I channel than the IA scheme, the reciprocity between the HKIA schemes for the Type I and Type II channels is weak, as can be seen by comparing Fig. 17 to Fig. 15. Here  $[1, 1, 1, 1]_{\text{I}}$  and  $[2, 2, 2, 2]_{\text{I}}$  are the HKIA schemes on the  $5 \times 6$  channel with beamforming reciprocal to that for the  $[1, 1, 1, 1]_{\text{II}}$  and  $[2, 2, 2, 2]_{\text{II}}$ , respectively, on the  $6 \times 5$  channel earlier. Note that the  $[2, 2, 2, 2]_{\text{I}}$  HKIA again outperforms IA, showing the benefits of introducing public

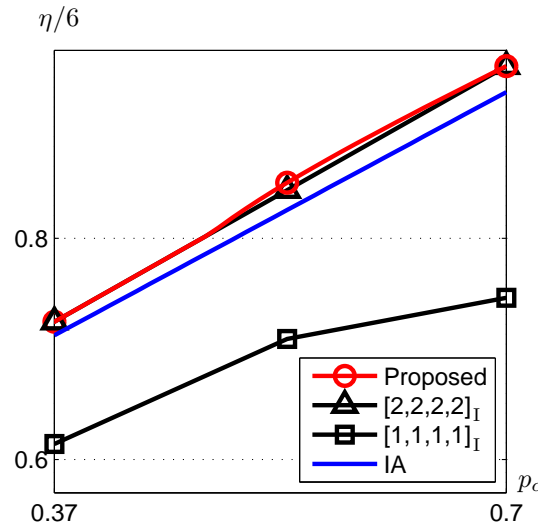


Fig. 17. Three HKIA schemes for the  $5 \times 6$  bursty channel with  $p_d = 0.7, p_{d|c} = 0.9, \frac{p_d}{1+p_{d|c}} < p_c \leq p_d$

messages, although the gain is not as great as before. However, somewhat surprisingly the  $[1, 1, 1, 1]_I$  performs strictly worse than the  $[2, 2, 2, 2]_I$ , and this is in fact true in general. Therefore, we base our HKIA scheme for the Type I channel on the  $[2, 2, 2, 2]_I$  HKIA whose beamforming as shown in Fig. 16(a), and add power tuning for private messages to it. (No blending of HKIA schemes here.) This leads to ‘Proposed’ curve in Fig. 17 on the  $5 \times 6$  channel. The non-reciprocity between the HKIA schemes on the Type I and Type II channels will be discussed further in Section VIII.

2) *General HKIA Scheme:* The general HKIA scheme for the Type I channel also closely parallels the one for the Type II channel in Section VI-D, with the only difference again being the encoding of the private messages. Specifically, when  $N$  is a multiple of 3, the  $[\frac{N}{3}, \frac{N}{3}, \frac{N}{3}, \frac{N}{3}]_I$  INBF scheme shown in Fig. 16(a) is employed together with the following random coding scheme:

$$X_1^n(M_{01}, M_{11}, M_{21}) = U_1^n(M_{01}) + \begin{bmatrix} G_1^{-1} \\ 0 \end{bmatrix} \begin{bmatrix} D_1^n(M_{11}) \\ C_1^n(M_{21}) \end{bmatrix}, \quad (24)$$

where  $U_1 \sim \mathcal{N}(0, \frac{P}{2M} I_M)$ ,  $C_1 \& D_1 \sim \mathcal{N}(0, \frac{P^a}{4N/3} I_{N/3})$ ,  $a \in [0, 1]$  and  $G_1 = \begin{bmatrix} \psi_1^* \hat{H}_{11} \\ \phi_2^* \hat{H}_{21} \end{bmatrix}$  is the effective channel matrix of the first broadcast channel (i.e. BC1 in Fig. 16(a)), with  $\hat{H}_{ji}$  denoting the first  $\frac{2}{3}N$  columns of  $H_{ji}$ .  $X_2^n(M_{02}, M_{12}, M_{22})$  is encoded symmetrically:

$$X_2^n(M_{02}, M_{12}, M_{22}) = U_2^n(M_{02}) + \begin{bmatrix} G_2^{-1} \\ 0 \end{bmatrix} \begin{bmatrix} D_2^n(M_{22}) \\ C_2^n(M_{12}) \end{bmatrix}, \quad (25)$$

where  $(U_2, C_2, D_2)$  have the same distributions as  $(U_1, C_1, D_1)$ , and  $G_2$  is the effective channel matrix of BC2. Appendix E extends this HKIA scheme for arbitrary  $N = 3k + q$ , where  $k$  is an integer and  $q \in \{0, 1, 2\}$ , leading to Proposition 3, which establishes  $\eta_{lb,1}$  of Theorem 3 and is proved in Appendix E.

*Proposition 3:* For the HKIA scheme in Section VII-C and Appendix E, the optimal value of ‘ $a$ ’ and the achievable sum DoF are as follows:

$$a_{\text{opt}} = \begin{cases} 1, & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_1 p_{d|c}} \\ \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - \lceil \frac{q}{2} \rceil, 2k + \lceil \frac{q}{2} \rceil \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_1 p_{d|c}}, \end{cases}$$

$$\eta_{\text{total}} \leq \begin{cases} \langle 4k + q, 2k + q, 2M - 2k - q \rangle, & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_1 p_{d|c}} \\ \langle N, M, M \rangle + \left\langle k, \left\lfloor \frac{q}{2} \right\rfloor, -\left\lfloor \frac{q}{2} \right\rfloor \right\rangle \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - \lceil \frac{q}{2} \rceil, 2k + \lceil \frac{q}{2} \rceil \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_1 p_{d|c}}, \end{cases}$$

where  $\beta_1 \triangleq \frac{2M - 4k - q}{2k + q}$ .

## VIII. DISCUSSION

We discuss some open and subtle issues in this section, starting with the non-reciprocity of the HKIA scheme on the Type I and the Type II channel, as observed in the previous section.

### A. The non-reciprocity of the HKIA scheme on the Type I and the Type II channel

The non-reciprocity of the HKIA schemes is not due to the fundamental BC-MAC difference as shown in Section IV-B. Rather, it is because introducing public messages is more effective on the Type II channel than on the Type I channel. To see this, let us look into the striking difference between the  $[1, 1, 1, 1]_{\text{II}}$  and the  $[1, 1, 1, 1]_{\text{I}}$  HKIA scheme in Fig. 15 and Fig. 17 for the  $6 \times 5$  and the  $5 \times 6$  channel. First of all, note that the INBF scheme of the  $[1, 1, 1, 1]_{\text{I}}$  and the  $[1, 1, 1, 1]_{\text{II}}$  HKIA as shown in Fig. 18 are clearly reciprocal, decomposing the channel into two orthogonal MACs and BC, respectively. Moreover, it is easily verified that they also achieve *identical* private DoF, i.e.  $2(p_c + p_d) = 2\langle 2, 1, 1 \rangle_{\mathcal{P}_S}$ !

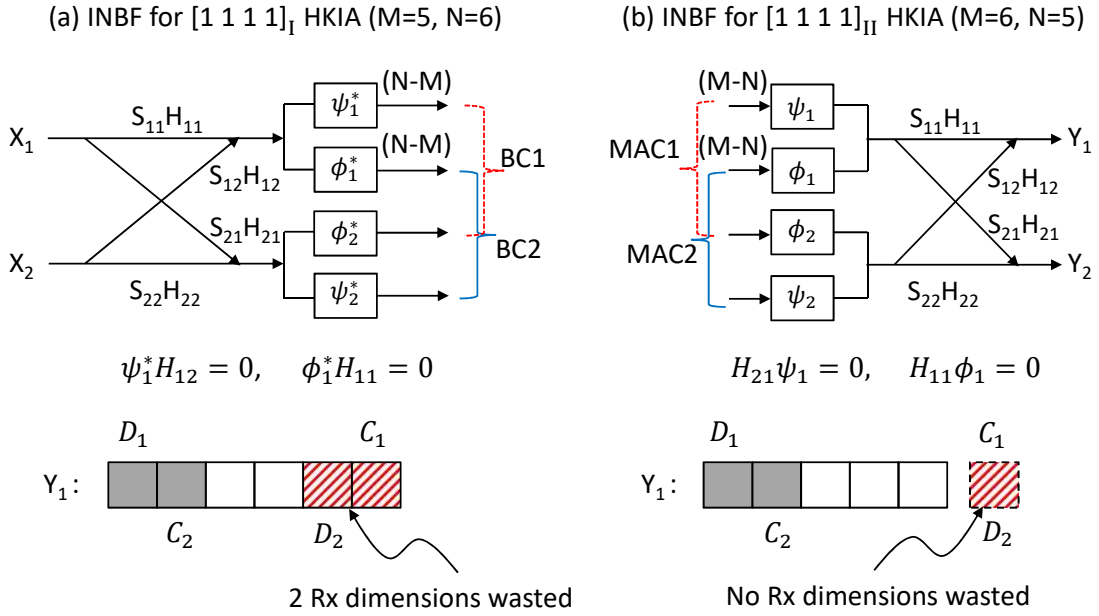


Fig. 18. The INBF efficiency for the  $[1, 1, 1, 1]_{\text{I}}$  and the  $[1, 1, 1, 1]_{\text{II}}$  HKIA scheme

However, the achievable public DoF is markedly different:  $\min(\langle 6, 0, 8 \rangle_{\mathcal{P}_S}, \langle 3, 4, 4 \rangle_{\mathcal{P}_S})$  is achievable by the  $[1, 1, 1, 1]_{\text{II}}$  but only  $\min(\langle 4, 0, 6 \rangle_{\mathcal{P}_S}, \langle 2, 3, 3 \rangle_{\mathcal{P}_S})$  is achieved by the  $[1, 1, 1, 1]_{\text{I}}$ . The reason for the low public DoF of the  $[1, 1, 1, 1]_{\text{I}}$  HKIA is actually obvious upon an inspection of the received private signal dimensions at each receiver. Recall that, from (12), (13), (24) and (25),  $D_1$  and  $C_2$  carry the messages to Rx1, while  $D_2$  and  $C_1$  convey the messages to Rx2. With the  $[1, 1, 1, 1]_{\text{II}}$  INBF, the desired signals  $D_1$  and  $C_2$  take two dimensions at Rx1, while undesired  $D_2$  and  $C_1$  are aligned and nulled, consuming no Rx dimensions, as illustrated in Fig. 18(b). On the other hand, since the  $[1, 1, 1, 1]_{\text{I}}$  INBF is done on the receivers, while desired  $D_1$  and  $C_2$  signals again occupy two Rx dimensions, the undesired  $D_2$  and  $C_1$  signals are not nulled at Rx1, when we decode the public messages. As a result,  $D_2$  and  $C_1$  present two extra dimensions of interference (noise) to the public messages at Rx1, thus impairing the achievable public DoF, as illustrated in Fig. 18(a). The situation at Rx2 is the same. Therefore, the INBF scheme works efficiently for the public message decoding on the Type II channel, but not so on the Type I

channel. This explains why the  $[1, 1, 1, 1]_{\text{II}}$  achieves higher DoF than the  $[1, 1, 1, 1]_{\text{I}}$ , and is even DoF-optimal on the Type II channel in part of Regime 2 as shown in Section VI-D.

The non-reciprocity of the HKIA schemes can also be seen by comparing the messaging schemes of the  $[1, 1, 1, 1]_{\text{I}}$  and the  $[1, 1, 1, 1]_{\text{II}}$  HKIA, as illustrated in Fig. 19, where we view the  $[1, 1, 1, 1]_{\text{I}}$  messages from a BC perspective and the  $[1, 1, 1, 1]_{\text{II}}$  messages from a MAC perspective. Recall that the  $[1, 1, 1, 1]_{\text{II}}$  HKIA decodes the public messages ( $M_{01}$  and  $M_{02}$ ) first as a MAC, and then decodes the private messages ( $M_{11}$  and  $M_{12}$ ) as another MAC. In terms of the BC-MAC duality, the  $[1, 1, 1, 1]_{\text{I}}$  HKIA is apparently not dual to the  $[1, 1, 1, 1]_{\text{II}}$ . However, from this perspective, it is also unclear how to make corrections to the  $[1, 1, 1, 1]_{\text{I}}$  HKIA to make it dual.

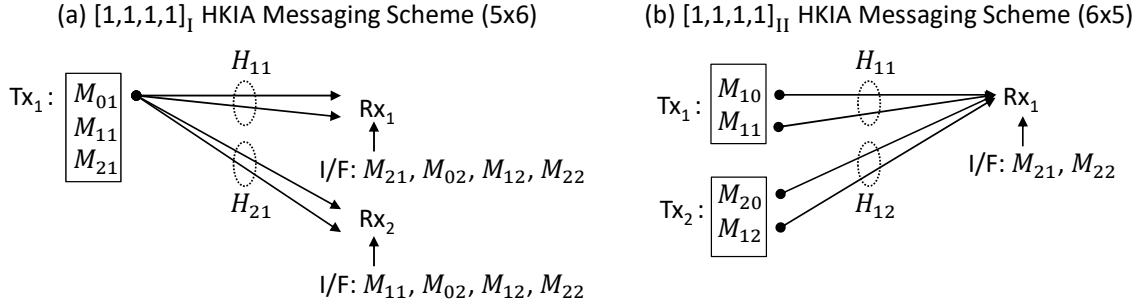


Fig. 19. Comparison of the  $[1, 1, 1, 1]_{\text{I}}$  and the  $[1, 1, 1, 1]_{\text{II}}$  HKIA messaging scheme (I/F: interference)

### B. The frequency hopping scenario

There is a subtlety that warrants some attention when we apply our results to applications where frequency hopping is allowed, as mentioned in Section I. Recall that our definition of DoF essentially measures the spectral efficiency of the system, c.f. (1). Since the MIMO X channel is spectrally more efficient than the respective point-to-point MIMO channel, having the two user pairs hop to two different frequency bands leads to lower spectral efficiency. So when there are two frequency bands available, we should employ cross-link messaging in both bands, instead of frequency hopping, if higher spectral efficiency is desired. However, an implicit premise of the above argument is the availability of multi-carrier capability in the system.

For applications or scenarios without multi-carrier capability, frequency hopping still provides a simple way to increase the sum rate of the two user pairs, e.g. in the earlier IEEE 802.11a/b/g wireless LANs. In this case, frequency hopping may be one of the reasons for the absence of the cross-links in our channel model, and we may lend our results to this scenario. Nonetheless, it should be understood that the DoF computed now is a special normalized sum rate over possibly two frequency bands, not the usual spectral efficiency on the two frequency bands. For example, suppose each transmitter and receiver in the system is designed for 20MHz bandwidth and we have a pool of 20MHz frequency bands for frequency hopping. Then the DoF ( $\eta$ ) signifies that, at high SNR, the system capacity is approximately the total capacity of  $\eta$  AWGN channels, each of 20MHz bandwidth, not 40MHz.

## IX. CONCLUSION

In this paper, we make progress on extending our knowledge of the sum DoF of the MIMO X channel, by considering a simple bursty channel model without feedback. The sum DoF of this bursty channel is fully characterized in Regime 1, and is partially characterized in Regime 2. The DoF characterization for  $\frac{2}{3} < \frac{M}{N} < \frac{\langle 3, -2, 2 \rangle_{\mathcal{P}_S}}{\langle 2, -1, 1 \rangle_{\mathcal{P}_S}}$  in Regime 2 remains a quest.

We also show striking differences between the bursty and the non-bursty MIMO X channel. In particular, burstiness of the channel creates new channel states, which may be under-utilized by the IA schemes. Further rate splitting and introducing public messages can achieve strictly higher DoF than the IA schemes, and is even DoF-optimal in some cases. Moreover, the Type I and the Type II channel are no longer reciprocal under this bursty channel model. And the DoF of the bursty X channel does not saturate when  $\frac{\min(M, N)}{\max(M, N)}$  reaches  $\frac{2}{3}$ .

While the IA schemes still play an important role on the bursty MIMO X channel, new coding schemes are needed to cope with the extra channel states. When  $p_c$  is small (Regime 1), treating crosslink interference as erasure/noise turns out to be DoF-optimal. As  $p_c$  grows (Regime 2), however, the paradigm shifts from avoiding crosslinks to exploiting them, and the usual IA schemes become useful. Furthermore, due to the fundamental difference between bursty BC and MAC with the burstiness known only to the receivers, the Type II channel needs more complex beamforming schemes than the Type I channel. The Type II channel even needs to overlay the IA signals with public messages as in HKIA to make full use of the channel states and achieve the DoF in Regime 2. When  $\frac{M}{N}$  is close to 1 in Regime 2, HKIA also proves to be beneficial on both the Type I and the Type II channel.

## REFERENCES

- [1] S. A. Jafar, and M. J. Fakhreddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, pp. 2637-2642, Jul. 2007.
- [2] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Signaling over MIMO multi-base systems: combination of multi-access and broadcast schemes," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2104-2108, Jul. 2006.
- [3] S. A. Jafar, and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 151-170, Jan. 2008.
- [4] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3457-3470, Aug. 2008.
- [5] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Communication through collisions: opportunistic utilization of past receptions," *Proc. IEEE INFOCOM*, pp. 2553-2561, Apr. 2014.
- [6] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Binary fading interference channel with no CSIT," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 666-670, Jul. 2014. Extended version: [arXiv:1405.0203v2](https://arxiv.org/abs/1405.0203v2)
- [7] I.-H. Wang, C. Suh, S. Diggavi, and P. Viswanath, "Bursty interference channel with feedback," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 21 - 25, July 2013. Extended version: <http://homepage.ntu.edu.tw/~Eihwang/Research/Eprints/ISIT13IC.pdf>
- [8] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vo. 27, pp. 49-60, Jan. 1981.
- [9] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534-5562, Dec. 2008.
- [10] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probability*, vol. 2, pp. 805-814, 1974
- [11] N. Khude, V. Prabhakaran, and P. Viswanath, "Opportunistic interference management," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2076-2080, Jul. 2009

- [12] S. Mishra, I.-H. Wang, and S. Diggavi, "Opportunistic interference management for multicarrier systems," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 389-393, Jun. 2013
- [13] S. Mishra, I.-H. Wang, and S. Diggavi, "Harnessing bursty interference in multicarrier systems with feedback," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 671-675, Jul. 2014
- [14] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, pp. 6093-6130, Oct. 2014.
- [15] S.-Y. Yeh, and I.-H. Wang, "Degrees of freedom of the bursty MIMO X channel without feedback," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1312-1316, Jul. 2016.
- [16] A. El Gamal, and Y.-H. Kim, *Network Information Theory*, Cambridge University Press, Cambridge, 2011.
- [17] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory* vol. 55, pp. 3032-3050, Jul. 2009
- [18] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory* vo. 33, pp. 711-716, Sep. 1987

## APPENDIX A

### PROOF OF SUM DoF UPPER BOUND FOR REGIME 2

For  $|\mathcal{I}_u^A| > |\mathcal{I}_u^B|$ , c.f. Fig. 11(b), (5) is similarly bounded by

$$\begin{aligned}
 & h(Y_1^n | S^n = u^n, M_{21}) - h((S_{21}H_{21}X_1 + Z_2)^n | S^n = v^n, M_{21}) \\
 & \leq h(\Omega_1 | \Omega_2) - h(Z_2^n) \\
 & \stackrel{(a)}{\leq} h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_v^B}\right) + h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A \setminus \mathcal{I}_v^B} | (H_{21}X_1 + Z_2)^{\mathcal{I}_u^A \setminus \mathcal{I}_v^B}\right) \\
 & \quad + h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_u^{BC}} | (H_{21}X_1 + Z_2)^{\mathcal{I}_u^{BC}}\right) + h\left((H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}\right), \tag{26}
 \end{aligned}$$

where (a) follows from the chain rule and the fact the conditioning reduces entropy. And again due to symmetry we also have

$$\begin{aligned}
 & h(Y_2^n | S^n = v^n, M_{12}) - h((S_{12}H_{12}X_2 + Z_1)^n | S^n = u^n, M_{12}) \\
 & \stackrel{(a)}{\leq} h\left((H_{21}X_1 + H_{22}X_2 + Z_2)^{\mathcal{I}_u^B}\right) + h\left((H_{21}X_1 + H_{22}X_2 + Z_2)^{\mathcal{I}_v^A \setminus \mathcal{I}_u^B} | (H_{12}X_2 + Z_1)^{\mathcal{I}_v^A \setminus \mathcal{I}_u^B}\right) \\
 & \quad + h\left((H_{22}X_2 + Z_2)^{\mathcal{I}_v^{BC}} | (H_{12}X_2 + Z_1)^{\mathcal{I}_v^{BC}}\right) + h\left((H_{21}X_1 + Z_2)^{\mathcal{I}_v^D}\right). \tag{27}
 \end{aligned}$$

Following the same single-letterization steps in Section V-D, each of the entropy terms in (26) can then be bounded by

$$\begin{aligned}
 & h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_v^B}\right) \leq |\mathcal{I}_v^B| h(H_{11}X_{1G_{\hat{\alpha}}} + H_{12}X_{2G_{\hat{\alpha}}} + Z_1), \\
 & h\left((H_{11}X_1 + H_{12}X_2 + Z_1)^{\mathcal{I}_u^A \setminus \mathcal{I}_v^B} | (H_{21}X_1 + Z_2)^{\mathcal{I}_u^A \setminus \mathcal{I}_v^B}\right) \\
 & \leq (|\mathcal{I}_u^A| - |\mathcal{I}_v^B|) h(H_{11}X_{1G_{\hat{\beta}}} + H_{12}X_{2G_{\hat{\beta}}} + Z_1 | H_{21}X_{1G_{\hat{\beta}}} + Z_2), \\
 & h\left((H_{11}X_1 + Z_1)^{\mathcal{I}_u^{BC}} | (H_{21}X_1 + Z_2)^{\mathcal{I}_u^{BC}}\right) \leq (|\mathcal{I}_u^B| + |\mathcal{I}_u^C|) h((H_{11}X_{1G_{\hat{\gamma}}} + Z_1) | (H_{21}X_{1G_{\hat{\gamma}}} + Z_2), \text{ and} \\
 & h\left((H_{12}X_2 + Z_1)^{\mathcal{I}_u^D}\right) \leq |\mathcal{I}_u^D| h((H_{12}X_{2G_{\hat{\delta}}} + Z_1), \tag{28}
 \end{aligned}$$

where  $X_{1G_{\hat{\alpha}}}$ ,  $X_{2G_{\hat{\alpha}}}$ ,  $X_{1G_{\hat{\beta}}}$ ,  $X_{2G_{\hat{\beta}}}$ ,  $X_{1G_{\hat{\gamma}}}$ , and  $X_{2G_{\hat{\delta}}}$  parallel the respective Gaussian random vectors in Section V-D.



Therefore, due to the law of large numbers and Fact 1, it follows that, for  $P_A > P_B$ , as  $n \rightarrow \infty$

$$\begin{aligned}
& \frac{1}{n} \mathcal{D}[h(Y_1^n | S^n, M_{21}) - h((SH_{21}X_1 + Z_2)^n | S^n, M_{21})] \\
& \leq P_B \min(2M, N) + (P_A - P_B) \min(M, N) + (P_B + P_C)[\min(M, 2N) - \min(M, N)] + P_D \min(M, N) \\
& = P_B \min(2M, N) + (P_B + P_C) \min(M, 2N) + (P_A - 2P_B) \min(M, N) \\
& = (p_d - p_c) \min(2M, N) + (p_d - p_{cd}) \min(M, 2N) + (p_{cd} - 2p_d + 2p_c) \min(M, N) \quad (\text{a.s.}).
\end{aligned} \tag{29}$$

Since (27) is symmetric to (26), in light of (4) we have proved Theorem 1 for  $p_c > \frac{p_d}{1+p_{d|c}}$ , which is equivalent to  $P_A > P_B$ .

## APPENDIX B

### THE RANK OF $(\theta_1^* H_{12} \phi_2)$ IN TYPE II, REGIME 1 BEAMFORMING

The fact that  $\theta_1^* H_{12} \phi_2$  is full-rank (a.s.) can be seen as follows. First of all, note that  $\theta_1^* H_{12}$  is full-rank (a.s.). Defining  $A \triangleq \theta_1^* H_{12}$ , it suffices to show that  $\mathcal{P}\{A\phi_2 \text{ is singular}\} = 0$ . Suppose  $\phi_2 = [u_1 \ u_2 \ \dots \ u_{M-N}]$ , where  $u_1, u_2, \dots, u_{M-N}$  are chosen randomly, independently, and uniformly within the unit ball of the null space of  $H_{22}$ . (Note that  $\phi_2$  is a function of  $H_{22}$ , independent of  $A$ , and is full-rank (a.s.).) Then  $\{A\phi_2 \text{ is singular}\}$  implies that some column of  $A\phi_2$  must be a linear combination of the rest columns. Without loss of generality, assume  $Au_1 = c_2 Au_2 + c_3 Au_3 + \dots + c_{M-N} Au_{M-N}$  for some  $c_2, c_3, \dots, c_{M-N}$ . Given  $A, u_2, u_3, \dots, u_{M-N}$ , this would constrain  $u_1$  to  $M - (M - N) + (M - N - 1) = M - 1$  dimensions, as there are  $(M - N)$  linear constraints and  $(M - N - 1)$  free variables. However,  $H_{22}$  consists of i.i.d. elements drawn from a continuous distribution, so by symmetry its null space (and hence  $u_1, u_2, \dots, u_{M-N}$ ) must spread over  $M$  dimensions. Moreover, given  $A, u_2, u_3, \dots$ , and  $u_{M-N}$ , it is not hard to see that  $u_1$  is still continuously distributed in the  $M$ -dimensional space. Therefore,  $\mathcal{P}\{u_1 \in (M - 1) \text{ dimensional subspace} \mid A, u_2, u_3, \dots, u_{M-N}\} = 0$ , and hence  $\mathcal{P}\{Au_1 \in \text{Col}[Au_2 \ Au_3 \ \dots \ Au_{M-N}]\} = 0$ . It then follows easily that  $\mathcal{P}\{A\phi_2 \text{ is singular}\} = 0$ .

## APPENDIX C

### PROOF OF (15)

*Fact 3:* If  $X$  is a Gaussian random vector distributed according to  $\mathcal{N}(0, (kP)I_M)$  for some positive constant  $k$ ,  $H$  is a  $N \times M$  matrix, and  $Z \sim \mathcal{N}(0, I_N)$ , then  $\mathcal{D}\{h(HX + Z)\} = \text{rank}(H)$ .

We show  $\mathcal{D}\{h(Y_1 | S_1, U_1)\} = \langle N, M - N, N \rangle$  (a.s.) here, using the simple fact above. Other terms in Section VI and other places can be proved along similar lines. First of all, letting  $V_1 \triangleq \psi_1 D_1 + \phi_1 C_1$  denote the random vector carrying the private messages at Tx1, we have

$$\begin{aligned}
& \mathcal{D}\{h(Y_1 | S_1, U_1)\} \\
& = \mathcal{D}\{p_{cd}h(H_{11}V_1 + H_{12}X_2 + Z_1) + p_{\bar{c}d}h(H_{11}V_1 + Z_1) + p_{c\bar{d}}h(H_{12}X_2 + Z_1) + (1 - p_c - p_d + p_{cd})h(Z_1)\} \\
& \stackrel{(a)}{=} \mathcal{D}\{p_{cd}h(H_{11}\psi_1 D_1 + H_{12}X_2 + Z_1) + p_{\bar{c}d}h(H_{11}\psi_1 D_1 + Z_1) + p_{c\bar{d}}h(H_{12}X_2 + Z_1)\}
\end{aligned} \tag{30}$$

where (a) follows because  $H_{11}\phi_1 = 0$  and the noise term can be dropped in DoF computation. Note that

$$H_{11}\psi_1 D_1 + H_{12}X_2 = H_{11}\psi_1 D_1 + H_{12}\phi_2 C_2 + H_{12}U_2 = [H_{11} \ H_{12}\phi_2 \ H_{12}] \begin{bmatrix} D_1 \\ C_2 \\ U_2 \end{bmatrix},$$

so we must have  $\mathcal{D}\{h(H_{11}\psi_1 D_1 + H_{12}X_2 + Z_1)\} \leq N$ . On the other hand, we are also lower bounded by  $\mathcal{D}\{h(H_{11}\psi_1 D_1 + H_{12}X_2 + Z_1)\} \geq \mathcal{D}\{h(H_{12}U_2 + Z_1)\} = N$  (a.s.). Hence we conclude that  $\mathcal{D}\{h(H_{11}\psi_1 D_1 + H_{12}X_2 + Z_1)\} = N$  (a.s.). For the same reason, obviously  $\mathcal{D}\{h(H_{12}X_2 + Z_1)\} = N$  (a.s.). Finally, it follows easily from Fact 3 that  $\mathcal{D}\{h(H_{11}\psi_1 D_1 + Z_1)\} = \text{rank}(H_{11}\psi_1) = M - N$  (a.s.). So the proof is completed.

#### APPENDIX D

##### PROOF OF PROPOSITION 2: HKIA LOWER BOUND FOR THE TYPE II CHANNEL

1) *Extending the HKIA scheme for arbitrary  $M$* : The key issue in extending the HKIA scheme in Section VII-B to the Type II channel with arbitrary  $M = 3k + q$  (i.e.  $k = \lfloor \frac{M}{3} \rfloor$  and  $q \in \{0, 1, 2\}$ ) is: What changes should be made to the INBF scheme in Fig. 16(b) when  $M/3$  is not an integer? Since the HKIA lower bound in Section VI-D is partially tight for  $N/M > 2/3$  in Regime 2, naturally we would continue to include its INBF scheme (Fig. 14(b)) as we did in Section VII-B. Similarly, the usual INBF scheme for the non-bursty channel [4] can be employed to enhance the lower bound when  $p_c$  becomes large, following the rationale in Section VII-B. To better appreciate these two HKIA schemes, we apply them on the  $7 \times 6$  channel and illustrate the achieved DoF in Fig. 20 as  $[1, 1, 1, 1]_{\text{II}}$  and  $[2, 2, 2, 3]_{\text{II}}$ . Note that the  $[2, 2, 2, 3]_{\text{II}}$  HKIA, which incorporates the usual INBF for the non-bursty channel, again outperforms IA and complements the  $[1, 1, 1, 1]_{\text{II}}$  HKIA when  $p_c$  is large.

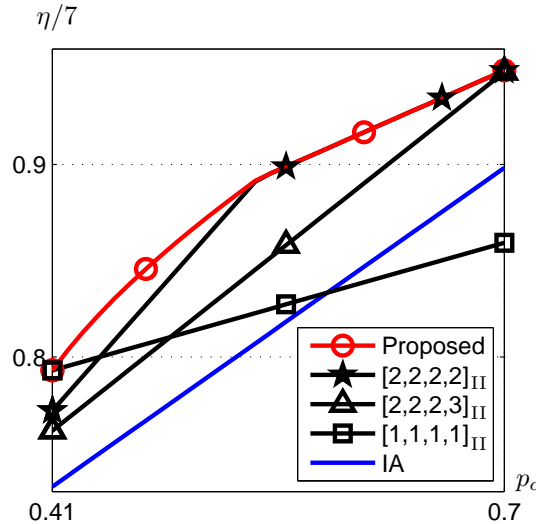


Fig. 20. HKIA schemes for the  $7 \times 6$  bursty channel with  $p_d = 0.7, p_{d|c} = 0.9, \frac{p_d}{1+p_{d|c}} < p_c \leq p_d$

Somewhat surprisingly, however, although the  $[2, 2, 2, 3]_{\text{II}}$  INBF scheme is the best known scheme for the non-bursty  $7 \times 6$  channel, it is not the best one to use here. To see this, consider an HKIA with the  $[2, 2, 2, 2]_{\text{II}}$  INBF which is similar to Fig. 16(b), except the dimensions of the  $\psi$  and  $\phi$  INBF matrices are now  $7 \times 2$  each, and the  $\hat{H}_{ij}$  represents 5 rows of the respective  $H_{ij}$  channel matrix. As shown in the figure, the  $[2, 2, 2, 2]_{\text{II}}$  HKIA achieves a strictly higher DoF than the  $[2, 2, 2, 3]_{\text{II}}$  HKIA, although the  $[2, 2, 2, 2]_{\text{II}}$  INBF is clearly suboptimal when the channel is non-bursty. As an astute reader may have thought about, this seemingly counter-intuitive result is in fact reasonable if we examine the utilization efficiency of received signal dimensions as we did in Section VIII-A, c.f. Fig. 18(b). Specifically, compared to the  $[2, 2, 2, 2]_{\text{II}}$  INBF, the  $[2, 2, 2, 3]_{\text{II}}$  is less efficient in utilizing the Rx signal dimensions because it consumes one more Rx dimension but that benefits only the Tx2-Rx2 link while causing more interference to the other three links. By freeing up this dimension and giving it to the public messages,  $[2, 2, 2, 2]_{\text{II}}$  HKIA is understandably more efficient. Our proposed HKIA scheme therefore blends the  $[2, 2, 2, 2]_{\text{II}}$  and  $[1, 1, 1, 1]_{\text{II}}$  INBF, and the achievable DoF on the  $7 \times 6$  channel is shown in Fig. 20.

The general HKIA scheme for arbitrary  $M$  blends the  $[(M - N), (M - N), (M - N), (M - N)]_{\text{II}}$  and the  $[k, k, k, k]_{\text{II}}$  INBF, similarly to what was done in Section VII-B. The  $[(M - N), (M - N), (M - N), (M - N)]_{\text{II}}$  INBF is just the one illustrated in Fig. 14(b), while the  $[k, k, k, k]_{\text{II}}$  INBF is similar to Fig. 16(b), except the  $\psi$  and  $\phi$  matrices are now  $M \times k$  matrices, and the  $\hat{H}_{ij}$  denotes  $M - k$  rows of  $H_{ij}$ . Specifically, the random coding in (12) and (13) now has

$$D_1, C_1, C_2, D_2 \sim \mathcal{N} \left( 0, \begin{bmatrix} \frac{1}{4(M-N)} P^b I_{M-N} & 0 \\ 0 & \frac{1}{4A} P^a I_A \end{bmatrix} \right) \quad (\text{i.i.d.}), \quad (31)$$

where  $A = k - (M - N)$ , and  $a, b \in [0, 1]$ . (It is easy to verify that  $A \geq 0$ .)

2) *Computation of the achievable sum DoF*: The analysis of the achievable DoF follows the same procedure for the HKIA scheme in Section VI-D3. First, the following private DoF is clearly achievable:

$$\eta_{\text{priv}} \leq 2g(a, b) \langle 2, 1, 1 \rangle, \quad (32)$$

where  $g(x, y) \triangleq xA + y(M - N)$  is a function introduced to simplify the expressions in this section. To compute the public DoF, we first re-evaluate (15) and (16), and it is not hard to verify that they now become

$$\begin{aligned} \mathcal{D}\{h(Y_1 | S_1)\} &= \langle N, N, N \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_1)\} &= \langle N, g(2a, b), N \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_2)\} &= \langle N, N, g(2a, b) \rangle, \text{ and} \\ \mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} &= \langle g(3a, 2b), g(2a, b), g(2a, b) \rangle, \end{aligned} \quad (33)$$

where  $g(3a, 2b) = g(2a, 2b) + g(a, 0)$  in the last equality.  $g(2a, 2b)$  is due to  $D_1$  and  $C_2$ , and  $g(a, 0)$  is due to the residual interference of  $C_1$  and  $D_2$  (aligned in  $A$  dimensions at  $Y_1$ .) By symmetry,  $\mathcal{D}\{h(Y_2 | S_2)\}$ ,  $\mathcal{D}\{h(Y_2 | S_2, U_2)\}$ ,  $\mathcal{D}\{h(Y_2 | S_2, U_1)\}$ , and  $\mathcal{D}\{h(Y_2 | S_2, U_1, U_2)\}$  can again be obtained readily. With these computed, (14)

in turn evaluates to

$$\begin{aligned}\eta_{01}, \eta_{02} &\leq \langle N, 0, N \rangle - \langle g(3a, 2b), 0, g(2a, b) \rangle, \text{ and} \\ \eta_{01} + \eta_{02} &\leq \langle N, N, N \rangle - \langle g(3a, 2b), g(2a, b), g(2a, b) \rangle.\end{aligned}\quad (34)$$

It then follows that

$$\eta_{\text{pub}} \leq \min \left( 2 \left[ \langle N, 0, N \rangle - \langle g(3a, 2b), 0, g(2a, b) \rangle \right], \quad \langle N, N, N \rangle - \langle g(3a, 2b), g(2a, b), g(2a, b) \rangle \right), \quad (35)$$

is achievable, and hence the total achievable DoF ( $\eta_{\text{total}} = \eta_{\text{priv}} + \eta_{\text{pub}}$ ) is

$$\eta_{\text{total}} \leq \min \left( 2 \left[ \langle N, 0, N \rangle + \langle g(-a, 0), g(a, b), g(-a, 0) \rangle \right], \quad \langle N, N, N \rangle + \langle g(a, 2b), g(0, b), g(0, b) \rangle \right). \quad (36)$$

3) *Optimization of ‘a’ and ‘b’*: Note that

$$\begin{aligned}\langle g(-a, 0), g(a, b), g(-a, 0) \rangle &= -aA\langle 1, -1, 1 \rangle + b(M - N)\langle 0, 1, 0 \rangle, \text{ and} \\ \langle g(a, 2b), g(0, b), g(0, b) \rangle &= aA\langle 1, 0, 0 \rangle + b(M - N)\langle 2, 1, 1 \rangle.\end{aligned}\quad (37)$$

Since the coefficients of ‘b’ is nonnegative in both terms, we should maximize ‘b’ and its optimal value is  $b = 1$ .

Substituting  $b = 1$  into (37) and (36), the total achievable DoF ( $\eta_{\text{total}}$ ) becomes

$$\min \left( 2 \left[ \langle N, M - N, N \rangle - aA\langle 1, -1, 1 \rangle \right], \quad \langle 2M - N, M, M \rangle + aA\langle 1, 0, 0 \rangle \right). \quad (38)$$

Recall that we are evaluating the DoF lower bound in Regime 2. From (38), it is straightforward to show that if  $\langle 3N - 2M, M - 2N, 2N - M \rangle_{\mathcal{P}_S} > 0$  or equivalently  $p_c/p_d > 1/(1 + \alpha p_{d|c})$ , we have

$$a_{\text{opt}} = \begin{cases} \frac{\langle 3N - 2M, M - 2N, 2N - M \rangle}{A \langle 3, -2, 2 \rangle}, & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_2 p_{d|c}}, \\ 1, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_2 p_{d|c}}, \end{cases} \quad (39)$$

where  $\beta_2 = \frac{q}{k+q}$ . Substituting (39) into (38) and simplifying the expressions, it is not hard to verify that

$$\eta_{\text{total}} \leq \begin{cases} M \left( \langle 1, 1, 1 \rangle + \langle 1, 0, 0 \rangle \frac{\langle 1, -1, 1 \rangle}{\langle 3, -2, 2 \rangle} \right), & \frac{p_c}{p_d} \leq \frac{1}{1 + \beta_2 p_{d|c}} \\ M \langle 1, 1, 1 \rangle + k \langle 1, 0, 0 \rangle, & \frac{p_c}{p_d} > \frac{1}{1 + \beta_2 p_{d|c}}, \end{cases} \quad (40)$$

is achievable, where we have used the identity  $\langle 1, 1, 1 \rangle + \langle 1, 0, 0 \rangle \frac{\langle 1, -1, 1 \rangle}{\langle 3, -2, 2 \rangle} = \langle 2, 1, 1 \rangle - \langle 1, 0, 0 \rangle \frac{\langle 2, -1, 1 \rangle}{\langle 3, -2, 2 \rangle}$ . Thus we have proved Proposition 2.

*Remark 5:* Since the following three inequalities are equivalent:

$$\langle 3N - 2M, M - 2N, 2N - M \rangle_{\mathcal{P}_S} > 0 \quad \equiv \quad \frac{p_c}{p_d} > \frac{1}{1 + \alpha p_{d|c}} \quad \equiv \quad \frac{N}{M} > \frac{\langle 2, -1, 1 \rangle_{\mathcal{P}_S}}{\langle 3, -2, 2 \rangle_{\mathcal{P}_S}},$$

the sum DoF lower bound in Proposition 2 can alternatively be viewed to hold when  $\frac{\langle 2, -1, 1 \rangle}{\langle 3, -2, 2 \rangle} < \frac{N}{M} \leq 1$  in Regime

2. (Note that  $\frac{2}{3} \leq \frac{\langle 2, -1, 1 \rangle}{\langle 3, -2, 2 \rangle}$ .)

*Remark 6:* Moreover, when  $\frac{p_c}{p_d} \leq \frac{1}{1 + \alpha p_{d|c}}$ , it is easily verified that the optimal value of  $a_{\text{opt}}$  is zero, and  $\eta_{\text{total}}$  reduces to  $2 \langle N, M - N, N \rangle$ , which is just the DoF achieved by the HKIA in Section VI-D2 when  $\frac{N}{M} > \frac{2}{3}$ . Therefore, the HKIA scheme for  $\frac{N}{M} > \frac{2}{3}$  in Section VI-D2 may be viewed as a special case of the scheme here.

## APPENDIX E

## PROOF OF PROPOSITION 3: HKIA LOWER BOUND FOR THE TYPE I CHANNEL

1) *Extending the HKIA scheme for arbitrary  $N$* : Not surprisingly, when  $M/N > 2/3$  and  $N$  is not a multiple of 3, the HKIA scheme for the Type I channel is not reciprocal to that for the Type II channel as described in Appendix D, due to the non-reciprocity of the HKIA scheme as discussed in Section VIII-A. For example, the  $[1, 1, 1, 1]_I$ ,  $[2, 2, 2, 3]_I$ , and  $[2, 2, 2, 2]_I$  HKIA DoF lower bound on the  $6 \times 7$  channel are illustrated in Fig. 21. It is evident in the figure that the  $[2, 2, 2, 3]_I$  HKIA, which incorporates usual INBF scheme for the non-bursty channel [4], delivers a strictly higher DoF than the other two HKIA schemes, again signifying the fact that introducing public messages on the Type I channel is not as effective as on the Type II channel. In this case, maximizing the dimensions used for orthogonal private message beamforming is more beneficial. Therefore we derive our lower bound by taking the  $[2, 2, 2, 3]_I$  HKIA scheme and fine-tuning the power of the private INBF signals as was done in Section VII-C. (No blending of HKIA schemes.) This proposed lower bound is also illustrated in Fig. 21.

For the general Type I channel with  $N = 3k + q$  where  $k$  is an integer and  $q \in \{0, 1, 2\}$ , the proposed HKIA scheme employs the  $[k, k, k, k]_I$ , the  $[k, k, k, (k+1)]_I$  or the  $[(k+1), k, k, (k+1)]_I$  INBF scheme, which is the usual INBF scheme for the non-bursty channel, when  $q = 0, 1$  or  $2$ , respectively. The decoding procedure and the analysis of the the achievable DoF again follow the lines in Section VI-D verbatim. We include it below for the sake of completeness. Note that it takes more effort to compute the achievable DoF on the Type I channel, as the form of the INBF scheme varies with  $q$  and can be asymmetric. We distinguish three cases.

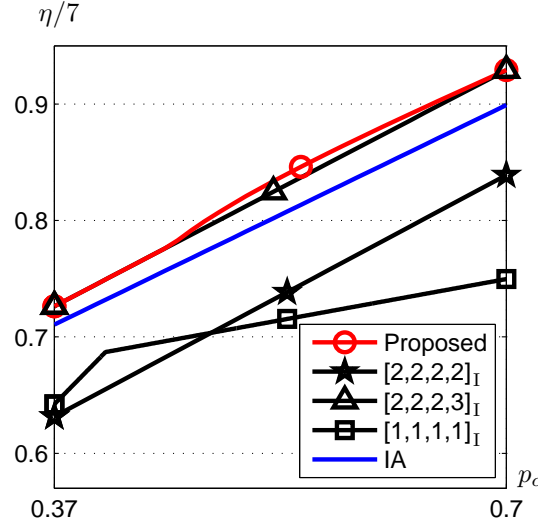


Fig. 21. HKIA schemes for the  $6 \times 7$  bursty channel with  $p_d = 0.7, p_{d|c} = 0.9, \frac{p_d}{1+p_{d|c}} < p_c \leq p_d$

2) *Proof of Proposition 3 when  $q = 0$ :* In this case, we have the  $\left[\frac{N}{3}, \frac{N}{3}, \frac{N}{3}, \frac{N}{3}\right]_I$  HKIA shown in Section VII-C2. (Note that  $N/3 = k$  when  $q = 0$ .) So the following private DoF is clearly achievable:

$$\eta_{\text{priv}} \leq 2ak(p_c + p_d) = ak \langle 4, 2, 2 \rangle. \quad (41)$$

As for the public DoF, with the new distributions of random variables, (15) can be verified to be

$$\begin{aligned} \mathcal{D}\{h(Y_1 | S_1)\} &= \langle N, M, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_1)\} &= \langle M + a(N - M), 2ak, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_2)\} &= \langle M + a(N - M), M, 2ak \rangle, \text{ and} \\ \mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} &= \langle aN, 2ak, 2ak \rangle. \end{aligned} \quad (42)$$

$\mathcal{D}\{h(Y_2 | S_2)\}$ ,  $\mathcal{D}\{h(Y_2 | S_2, U_2)\}$ ,  $\mathcal{D}\{h(Y_2 | S_2, U_1)\}$ , and  $\mathcal{D}\{h(Y_2 | S_2, U_1, U_2)\}$  are obtained immediately by symmetry. Hence, by straightforward computation, the achievable public DoF now becomes

$$\eta_{\text{pub}} \leq \min(2[\langle M, 0, M \rangle - a \langle M, 0, 2k \rangle], \quad \langle N, M, M \rangle - ak \langle 3, 2, 2 \rangle), \quad (43)$$

and the total achievable DoF is

$$\eta_{\text{total}} \leq \min(2[\langle M, 0, M \rangle - a \langle M - 2k, -k, k \rangle], \quad \langle N, M, M \rangle + ak \langle 1, 0, 0 \rangle). \quad (44)$$

Now the remaining task is to optimize ‘ $a$ ’ so that  $\eta_{\text{total}}$  is maximized. We distinguish two cases: If  $\langle M - 2k, -k, k \rangle \leq 0$ , then clearly we should maximize ‘ $a$ ’, i.e.  $a_{\text{opt}} = 1$ . Otherwise, the optimal ‘ $a$ ’ is obtained by equating  $2[\langle M, 0, M \rangle - a \langle M - 2k, -k, k \rangle]$  and  $\langle N, M, M \rangle + ak \langle 1, 0, 0 \rangle$ . Therefore, noting that  $\langle M - 2k, -k, k \rangle \leq 0$  is equivalent to  $p_c/p_d \leq 1/(1 + \frac{M-2k}{k}p_{d|c})$ , we have

$$a_{\text{opt}} = \begin{cases} 1, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left(\frac{M-2k}{k}\right)p_{d|c}} \\ \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k, 2k \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left(\frac{M-2k}{k}\right)p_{d|c}} \end{cases} \quad (45)$$

Substituting (45) into (44) and simplifying the expressions, we then conclude that the following total DoF is achievable:

$$\eta_{\text{total}} \leq \begin{cases} 2 \langle 2k, k, M - k \rangle, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left(\frac{M-2k}{k}\right)p_{d|c}} \\ \langle N, M, M \rangle + \langle k, 0, 0 \rangle \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k, 2k \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left(\frac{M-2k}{k}\right)p_{d|c}}. \end{cases} \quad (46)$$

3) *Proof of Proposition 3 when  $q = 1$ :* When  $q = 1$ , the usual  $[k, k, k, (k+1)]_I$  INBF for the non-bursty channel [4] is used in the HKIA, so the  $D_1, C_1, C_2$ , and  $D_2$  random vectors in (24) and (25) now have the following distribution

$$\begin{aligned} D_1, C_1, C_2 &\sim \mathcal{N}\left(0, \frac{P^a}{4k} I_k\right), \text{ and} \\ D_2 &\sim \mathcal{N}\left(0, \frac{P^a}{4(k+1)} I_{k+1}\right), \end{aligned} \quad (47)$$

where  $a \in [0, 1]$ . The achievable private DoF is hence easily seen to be

$$\eta_{\text{priv}} \leq a \langle 4k + 1, 2k + 1, 2k \rangle. \quad (48)$$

The determination of the achievable public DoF takes a little more effort due to the asymmetry between  $D_1$  and  $D_2$ , but it still parallels the steps in Section VII-C. So again the following differential entropies of  $Y_1$  are evaluated:

$$\begin{aligned} \mathcal{D}\{h(Y_1 | S_1)\} &= \langle N, M, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_1)\} &= \langle M + (N - M)a, 2ka, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_2)\} &= \langle M + (N - M)a, M, (2k + 1)a \rangle, \text{ and} \\ \mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} &= \langle Na, 2ka, (2k + 1)a \rangle. \end{aligned} \quad (49)$$

The differential entropies of  $Y_2$  is not symmetric to those of  $Y_1$  but can be verified to be

$$\begin{aligned} \mathcal{D}\{h(Y_2 | S_2)\} &= \langle N, M, M \rangle, \\ \mathcal{D}\{h(Y_2 | S_2, U_2)\} &= \langle M + (N - M)a, (2k + 1)a, M \rangle, \\ \mathcal{D}\{h(Y_2 | S_2, U_1)\} &= \langle M + (N - M)a, M, 2ka \rangle, \text{ and} \\ \mathcal{D}\{h(Y_2 | S_2, U_1, U_2)\} &= \langle Na, (2k + 1)a, 2ka \rangle. \end{aligned} \quad (50)$$

From (49) and (50), it is straightforward to verify that the achievable public DoF is

$$\eta_{\text{pub}} \leq \min \left( \langle 2M, 0, 2M \rangle - a \langle 2M, 0, 4k + 1 \rangle, \quad \langle N, M, M \rangle - a \langle N, 2k + 1, 2k \rangle \right), \quad (51)$$

It follows from (48) and (51) that the total achievable DoF is

$$\eta_{\text{total}} \leq \min \left( \langle 2M, 0, 2M \rangle - a \langle 2M - 4k - 1, -2k - 1, 2k + 1 \rangle, \quad \langle N, M, M \rangle + a \langle k, 0, 0 \rangle \right). \quad (52)$$

The optimal value of ‘ $a$ ’ can then be shown to be

$$a_{\text{opt}} = \begin{cases} 1, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left( \frac{2M - 4k - 1}{2k + 1} \right) p_{d|c}}, \\ \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - 1, 2k + 1 \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left( \frac{2M - 4k - 1}{2k + 1} \right) p_{d|c}}, \end{cases} \quad (53)$$

which leads to the following achievable DoF:

$$\eta_{\text{total}} \leq \begin{cases} \langle 4k + 1, 2k + 1, 2M - 2k - 1 \rangle, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left( \frac{2M - 4k - 1}{2k + 1} \right) p_{d|c}}, \\ \langle N, M, M \rangle + \langle k, 0, 0 \rangle \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - 1, 2k + 1 \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left( \frac{2M - 4k - 1}{2k + 1} \right) p_{d|c}}. \end{cases} \quad (54)$$

4) *Proof of Proposition 3 when  $q = 2$ :* When  $q = 2$ , the HKIA scheme incorporates the usual  $[(k+1), k, k, (k+1)]_I$  INBF for the non-bursty channel [4]. In this case,  $D_1, C_1, C_2$ , and  $D_2$  random vectors in (24) and (25) now take the following distribution

$$\begin{aligned} C_1, C_2 &\sim \mathcal{N}\left(0, \frac{P^a}{4k} I_k\right), \text{ and} \\ D_1, D_2 &\sim \mathcal{N}\left(0, \frac{P^a}{4(k+1)} I_{k+1}\right), \end{aligned} \quad (55)$$

where  $a \in [0, 1]$ . The achievable private DoF is hence

$$\eta_{\text{priv}} \leq 2a \langle 2k+1, k+1, k \rangle. \quad (56)$$

To evaluate the achievable public DoF, we first compute

$$\begin{aligned} \mathcal{D}\{h(Y_1 | S_1)\} &= \langle N, M, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_1)\} &= \langle M + (N - M)a, (2k+1)a, M \rangle, \\ \mathcal{D}\{h(Y_1 | S_1, U_2)\} &= \langle M + (N - M)a, M, (2k+1)a \rangle, \text{ and} \\ \mathcal{D}\{h(Y_1 | S_1, U_1, U_2)\} &= \langle Na, (2k+1)a, (2k+1)a \rangle. \end{aligned} \quad (57)$$

The differential entropies of  $Y_2$  is obtained readily due to symmetry. Therefore, the achievable public DoF is

$$\eta_{\text{pub}} \leq \min \left( 2 \langle M, 0, M \rangle - 2a \langle M, 0, 2k+1 \rangle, \quad \langle N, M, M \rangle - a \langle N, 2k+1, 2k+1 \rangle \right), \quad (58)$$

and the total achievable DoF is

$$\eta_{\text{total}} \leq \min \left( 2 \langle M, 0, M \rangle - 2a \langle M - 2k - 1, -k - 1, k + 1 \rangle, \quad \langle N, M, M \rangle + a \langle k, 1, -1 \rangle \right). \quad (59)$$

Straightforward optimization of ‘ $a$ ’ then leads to

$$a_{\text{opt}} = \begin{cases} 1, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left( \frac{2M-4k-2}{2k+2} \right) p_{d|c}}, \\ \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - 1, 2k + 1 \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left( \frac{2M-4k-2}{2k+2} \right) p_{d|c}}, \end{cases} \quad (60)$$

and the achievable DoF is

$$\eta_{\text{total}} \leq \begin{cases} \langle 4k+2, 2k+2, 2M-2k-2 \rangle, & \frac{p_c}{p_d} \leq \frac{1}{1 + \left( \frac{2M-4k-2}{2k+2} \right) p_{d|c}}, \\ \langle N, M, M \rangle + \langle k, 1, -1 \rangle \frac{\langle 2M - N, -M, M \rangle}{\langle 2M - N, -2k - 1, 2k + 1 \rangle}, & \frac{p_c}{p_d} > \frac{1}{1 + \left( \frac{2M-4k-2}{2k+2} \right) p_{d|c}}. \end{cases} \quad (61)$$

Finally, noting that (45), (46), (53), (54), (60) and (61) are equivalent to Proposition 3, we have completed its proof.